Exchange rates and their effect on international trade

Klaassen, F.J.G.M.

Publication date:
2000

CentER

Exchange Rates and Their Effects on International Trade

Franc J.G.M. Klaassen

Exchange Rates

and

Their Effects on International Trade

Exchange Rates

and

Their Effects on International Trade

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Katholieke Universiteit Brabant, op gezag van de rector magnificus, prof. dr. F.A. van der Duyn Schouten, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de aula van de Universiteit op

vrijdag 28 januari 2000 om 14.15 uur

door

FRANCISCUS JOHANNES GERARDUS MARIA KLAASSEN

geboren op 25 maart 1971 te Weert

PROMOTOR: Prof. dr. H.P. Huizinga

COPROMOTOR: Dr. F.C.J.M. de Jong

Voor mijn ouders

Acknowledgements

This dissertation is the result of the Ph.D. program I followed at CentER and the Departments of Econometrics and Economics of Tilburg University. I thank all people who stimulated me to write this book during the past four years.

I owe special gratitude to my promotor Harry Huizinga, copromotor Frank de Jong and to Siem Jan Koopman, who took over most of Frank's role when Frank left Tilburg University one year ago. I appreciate the mix of the research freedom they gave me and their excellent advice, which has undoubtedly improved the quality of this thesis.

Second, I thank Sylvester Eijffinger, Henk Jager and Theo Nijman for taking place in my dissertation committee, and also for giving valuable comments during my Ph.D. project.

There are several other people who have contributed to my work. In particular, I am grateful to Geert Bekaert, Herman Bierens, Herman van Dijk, Bas Donkers, Bertrand Melenberg, Arthur van Soest, and Casper de Vries. I also thank Jan Magnus, from whom I have learned a lot qua research attitude. I thank the Netherlands Network of Economics (NAKE) for organizing Ph.D. courses and workshops.

An important factor in my work was the sociable atmosphere at Tilburg University. In this respect, I want to thank Bas Donkers, with whom I have shared the office for almost four years. I am also grateful to all other colleagues, particularly my sports and dinner mates.

Last but not least, I am very grateful to my parents. Their support was crucial for my work on this dissertation.

Franc Klaassen
September 1999

Contents

Chapter 1

Introduction

Exchange rates and international goods trade are two important elements of the current international economic system. They have attracted many economic researchers. Yet, there are still numerous open questions concerning the processes behind exchange rates and trade.

This book addresses five of these issues from an empirical point of view. In short, we discuss the following topics (see below for a more detailed description). First, we analyze the volatility of exchange rates and provide a method to improve existing volatility forecasts. Then we test whether exchange rates exhibit long swings. Such swings may originate from changes in economic policy. For instance, the Volcker switch to an anti-inflationary policy in the United States in 1979 was probably one of the reasons behind the long U.S. dollar appreciation in the first half of the eighties. Third, we analyze the validity of the purchasing power parity (PPP) hypothesis, a building block of many economic theories: is the exchange rate proportional to the ratio of the two countries' price levels in the long-run? Next, we examine the correlations between exchange rates and their variation over time. For instance, did they rise and fall after the introduction and demise, respectively, of exchange rate coordination within the European Monetary System (EMS)? The fifth and final topic of this thesis concerns the supposedly negative effect of exchange rate risk on international trade. This effect plays a prominent role in discussions on the choice between a fixed and floating exchange rate regime and on the desirability of foreign exchange interventions to stabilize exchange rates.

For all five discussions we employ modern econometric techniques, either refined existing methods or newly developed techniques. In short, the following methods are discussed (a more extensive description follows below). We propose a new way of combining Markov regime switches in the variance with generalized autoregressive conditional heteroskedasticity (GARCH) processes in order to improve the standard GARCH volatility forecasts. We use models with regime switches in the mean instead of variance to capture long swings in exchange rates and to test for PPP. Moreover, we introduce a new multivariate GARCH model to describe exchange rate correlations over time. Finally, we show how daily exchange rates can help reduce measurement error in multi-months-ahead exchange rate risk measures that are relevant for goods traders.

In section 1.1 of this introductory chapter we motivate our work more extensively. The second section describes each of the five research topics in more detail and gives the basic ideas behind the econometric techniques that we use.

1.1 Motivation

The role of (spot) exchange rates and international (goods) trade has grown substantially over the last decades. For instance, the global foreign exchange market turnover for spot transactions has increased by 28 per cent in real terms over the last ten years and is about 600 billion U.S. dollars per day nowadays.[1] Furthermore, world goods exports have risen by 149 per cent in volume over the last twenty years, which is about three times the 52 per cent increase in the volume of goods production; exports have reached a value of 5303 billion U.S. dollars.[2]

It is not surprising that the importance of exchange rates and trade has increased simultaneously. After all, for most international trade transactions at least one of the trading partners has to deal with a foreign currency, so that there is often a direct link between foreign exchange transactions and trade flows.

This connection can have great implications, as we saw in the recent East-

[1] See Bank for International Settlements (1999); the real increase is based on the nominal increase of 71 per cent from April 1989 to April 1998 corrected for U.S. consumer price inflation and exchange rate changes.

[2] World Trade Organization (1998), in particular the figures for 1978 and 1997.

Asian crisis. Thailand devalued its currency, the baht, to stimulate exports and thereby the economy. This devaluation, however, lowered the competitiveness of its trading partners. Eventually, many currencies in the region devalued. We saw in Indonesia that the resulting price inflation can seriously disrupt the real economy and lead to social unrest.

Their importance and strong linkage have made exchange rates and trade the subject of many policy discussions. In late 1984, for instance, the U.S. dollar had become so expensive against the currencies of the main U.S. trading partners that several U.S. sectors demanded protective legislation. There was a threat that this would lead to a global chain of trade policy measures, as countries tend to respond by retaliation measures. To avoid this, the Group of Five (G-5) countries decided to try to bring the dollar down (Plaza agreement in 1985, see Krugman and Obstfeld (1991)). This policy intervention in the foreign exchange market was apparently successful, as the dollar strongly depreciated in the two years after that.

A second example of a policy discussion related to exchange rates and trade concerns European monetary unification (EMU). It is commonly assumed that the elimination of exchange rate risk benefits trade. The standard argument is that less exchange risk decreases the riskiness of trade profits, leading risk averse traders to increase trade. Hence, the relation between exchange rates and trade is one of the main economic arguments put forward in the debate on the desirability of EMU (EU Commission (1990)).

Given their relevance for policy issues, exchange rates and international trade have been the focus variables in many economic studies, both theoretical and empirical. Nevertheless, there is still no full understanding of their determinants and the mechanisms through which these determinants affect exchange rates and trade. For example, in the short run the role of fundamentals in the exchange rate generating process seems limited. Taylor and Allen (1992), for instance, find in their survey among foreign exchange dealers that the majority of them use chart analysis in forming exchange rate expectations up to, say, one week. As the horizon is lengthened, foreign exchange dealers give more weight to fundamental analysis. However, it is still an unresolved question what variables are truly relevant for exchange rate determination.

One open question concerning the determinants of exchange rates is the relevance of policy interventions. Policy shifts such as the Plaza announcement may affect the exchange rate trend for some time. This can lead to long swings in exchange rates. The question is whether such swings are indeed present in exchange rate data.

Another unresolved question is to what extent exchange rates are governed by price levels. For example, does relative purchasing power parity (PPP) hold in the long run. Many economists intuitively argue that PPP is true; goods arbitrage would equalize prices expressed in the same currency across countries. However, the existing empirical evidence is much less supportive (Froot and Rogoff (1996)).

A third question that is still under debate concerns the economic argument for EMU mentioned above: does exchange rate stability really benefit trade? Many economists intuitively think it does and this view is supported by regular opinion surveys of business leaders, such as those conducted by the Confederation of British Industry (1989). Nevertheless, the results in the empirical literature are ambiguous (Côté (1994)).

Unresolved economic questions such as the three just described provide the motivation for this book. It contributes to the literature on exchange rates and international trade in two aspects. First, we try to answer a number of significant economic questions, including the ones just presented. We do this from an empirical point of view, using the United States as the central country. Second, we often encounter that existing econometric techniques are not directly applicable to the issues we are interested in. Hence, we refine several techniques and, when necessary, introduce alternative approaches; these methodological contributions may be helpful in subsequent research. Therefore, this book can be positioned in the intersection of international economics, international finance and econometrics.

1.2 Overview

The thesis consists of five chapters, apart from this introduction and the concluding chapter.[3] Chapters 2 to 5 deal with exchange rates, whereas Chapter 6 is about the effect of exchange rates on trade. The chapters are self-contained. We now briefly describe each chapter in terms of focus, motivation, main existing

[3]Earlier versions of these five chapters have appeared as CentER Discussion Papers, namely Chapter 2 as No. 9852, 3 as 9908, 4 as 9909, 5 as 9910 and 6 as 9973.

studies and the results. This will clarify our contribution to the literature as given above.

Chapter 2: Improving GARCH Volatility Forecasts with a New Regime-Switching GARCH Model

In Chapter 2 we focus on exchange rate volatility. This is an important aspect of many financial decisions. For example, volatility of exchange rates is a determinant for pricing currency options. It may also affect international goods trade, as explained in section 1.1. Hence, there is a need for good volatility forecasts.

Many authors use forecasts based on generalized autoregressive conditional heteroskedasticity (GARCH) models. These models assume that current volatility is a function of the previous exchange rate surprise and past volatility (see Bollerslev, Chou and Kroner (1992) for an overview). The quality of such forecasts is stressed by Andersen and Bollerslev (1998).

We show that GARCH forecasts are, nevertheless, too variable and introduce a generalization of GARCH, regime-switching GARCH, that performs better in this respect.

The origin of the excess variability of GARCH forecasts may be the well-known high persistence of individual shocks in GARCH volatility forecasts. After all, if large shocks increase subsequent volatility forecasts for a long time, the variability of the forecasts rises.

It is not clear why all shocks should be persistent, as GARCH usually implies. Shocks may be "pressure relieving" instead of persistent, so that they are followed by a period of low instead of high volatility. The relevance of this effect is illustrated by the period after the Plaza announcement in 1985. The day after the agreement, the dollar depreciated strongly. However, in the subsequent months, the foreign exchange market was relatively quiet. The sharp fall may have relieved the foreign exchange market from the tensions that had resulted from the sharp dollar appreciation in the years before.

To improve on the GARCH volatility forecasts we want to allow for more flexibility regarding volatility persistence of shocks. To this end, we introduce an additional source of volatility persistence based on the Hamilton (1989) Markov regime-switching model. That is, we allow for two regimes with different volatility levels. Persistence of these regimes already creates volatility persistence. In addition, we allow for two different GARCH-type models to govern volatility within the regimes. Hence, shocks can persist due to regime persistence as well as GARCH effects. They can also be pressure relieving, if the shock is followed by a switch to the low volatility or low persistence regime. This creates extra flexibility compared to standard, single-regime GARCH.

The way we combine the regime-switching model with GARCH into a regime-switching GARCH model is novel. The most important advantage over existing variants, such as Gray (1996a), is that multi-period-ahead volatility forecasting becomes a convenient procedure. This enable us to compare standard, single-regime GARCH with regime-switching GARCH forecasts. The empirical study on the three main U.S. dollar exchange rates (German mark, Japanese yen, U.K. pound) shows that regime-switching GARCH (out-of-sample) forecasts do not suffer from excess variability and outperform standard GARCH forecasts.

Chapter 3: Long Swings in Exchange Rates: Are They Really in the Data?

One of the issues that has received much attention in the literature so far is the process underlying the level of exchange rates. Many structural exchange rate models have been developed, but their empirical validity is often questioned, particularly in the short-run (see MacDonald and Taylor (1992) for an overview). Therefore, since Meese and Rogoff (1983), it is a widespread view that exchange rates follow a random walk. This has important consequences for exchange rate forecasting, as it implies that the best predictor of future exchange rates is the current exchange rate, possibly adjusted by a constant. The empirical quality of this simple forecasting rule has also been stressed by Diebold and Nason (1990).

The random walk, however, is unsatisfactory from an economic point of view. It ignores any effect of observed changes in economic policy, and, according to the Lucas (1976) critique, such policy shifts may well affect the exchange rate generating process. For instance, regarding monetary policy, the Volcker switch to a contractionary monetary policy in 1979 may well have increased the structural exchange rate appreciation, causing the strong dollar appreciation in the first half of the eighties. Moreover, the potential relevance of international policy coordination for the exchange rate process appears from the 1985 Plaza intervention and the subsequent strong dollar depreciation from 1985 to 1987. Both examples show that policy shifts can lead to changes in the trend of exchange rates and thus to long swings.

The focus of Chapter 3 is whether long swings exist. This issue is relevant for various reasons. First, if swings exist, this may be an indication of the relevance of economic policy for exchange rates, despite the empirical rejections of existing structural exchange rate models. Hence, the existence of long swing is important for future research on exchange rate determination.

Another reason to analyze the existence of long swings is that such swings can provide an explanation for peso problems. In other words, the existence of long swings can explain that exchange rate expectations of rational investors appear biased ex post for a long time. After all, if swings exist, rational investors incorporate the possibility of a swing reversal in their expectations, even though the swing reversal may not materialize in the actual exchange rate process for a long time. This leads to a long period of ex post biased expectations.

Econometrically, long swings can be modeled by the Markov regime-switching model described above, but with the crucial difference that the regimes now concern the mean instead of the variance. For instance, persistence of an appreciation and a depreciation regime can generate the dollar swing in the eighties. In Chapter 3 we formally examine the existence of long swings by testing the random walk against the more general regime-switching model, which is commonly used to model long swings (see Engel and Hamilton (1990), among others). We use data on the three main U.S. dollar exchange rates mentioned above.

Earlier studies such as Engel and Hamilton (1990) conclude that long swings exist. However, the authors are concerned about the reliability of their Wald based tests in the strongly nonlinear regime-switching model. We show that their tests are indeed not very robust. Hence, we use the more robust likelihood ratio test. Remarkably, with similar quarterly data as Engel and Hamilton (1990) use, likelihood ratio tests yield no evidence of swings.

The lack of evidence from quarterly data may be due to the low data frequency: even if swings exist and last for some quarters, sampling at the quarterly frequency may result in too few observations per swing to distinguish the swings from a random walk. Therefore, we use monthly and weekly data to enhance the power of the test. Only the likelihood ratios for weekly data are significant. Hence, we eventually conclude that long swings are in the data.

Chapter 4: Purchasing Power Parity: Evidence from a New Test

Chapter 4 concerns one of the oldest theories in international economics, namely Purchasing Power Parity (PPP). As usual, we concentrate on long-run relative PPP, that is, the long-run proportionality of the exchange rate and the ratio of the two countries' price levels.

The empirical validity of PPP is an important issue in several respects. First, PPP is a building block of many traditional structural models of exchange rate determination, so that their validity depends on the validity of PPP. Second, knowing whether PPP holds is important for the development of new structural exchange rate models. From a practical point of view, PPP can also provide a target exchange rate for monetary authorities, which they can use for foreign exchange interventions, among other things. Finally, the long-term behavior of exchange rates relative to prices is relevant for international firms. They have to decide upon foreign direct investments and therefore require reliable forecasts of the real value of the long-lasting income stream generated by the investment projects. Taking account of long-run PPP, if valid, may help improve the long-run exchange rate forecasts they need.

Most economists intuitively consider this hypothesis to be true and use PPP as a building block in other theories. Quite surprisingly, however, the existing empirical literature is not very supportive of PPP (see Froot and Rogoff (1996)). In Chapter 4 we re-examine the empirical validity of PPP with a new test approach.

Most studies on PPP test whether the real exchange rate follows a random walk against the alternative of stationarity, that is, PPP. However, the previous chapter has presented evidence of long swings in nominal (and hence real) exchange rates. Therefore, it seems more appropriate to test PPP in a model that allows for such swings instead of a random walk context. As in Chapter 3, we use the Markov regime-switching mean model as the baseline model for the nominal exchange rate.

The next question is how to test for PPP in the long swings model. We show that PPP holds if a swing is likely to end when the PPP disequilibrium becomes large and if the next swing governs the exchange rate back to its PPP equilibrium. To test for PPP we thus examine whether these conditions are valid. This represents a new test approach for PPP.

Remarkably, our test yields evidence in favor of PPP in all three of the world's most important exchange rates mentioned above. This result supports exchange rate theories that incorporate PPP. It can also improve long-run exchange rate forecasts needed in practice. Indeed, we show that predictions of the direction of exchange rate changes improve when PPP is accounted for.

Given the evidence in favor of PPP, it is natural to examine what the economic mechanism behind PPP is. The common argument for PPP is that goods arbitrage equalizes prices in the same currency across countries. Because it is commonly believed that goods markets have become more integrated, making arbitrage easier, it is interesting to examine whether PPP disequilibria have become shorter-lived. We conclude that they have for the German mark and the U.K. pound, but not for the Japanese yen. This may indeed be explained by changes in trade openness, as we find that both European economies have become much more open, while Japan is still relatively closed.

Chapter 5: Have Exchange Rates Become More Closely Tied? Evidence from a New Multivariate GARCH Model

In Chapter 5 we leave the univariate regime-switching settings of the three previous chapters and focus on correlations between exchange rates and their development over time. Correlations are a key determinant of many financial decisions. For instance, investors in stocks need correlation assessments to reduce the riskiness of their portfolios. Correlations between exchange rates are also important for internationally trading corporations and banks, as they have to hedge open foreign exchange positions.

One of the few studies that deals with modeling exchange rate correlations is Bollerslev (1990), which analyzes correlations between several EMS (European Monetary System) - U.S. dollar exchange rates. Bollerslev develops a multivariate GARCH model in which he assumes that the correlations are constant over time.

Economic intuition, however, questions this constancy of exchange rate correlations. For instance, suppose the U.K. joins the Exchange Rate Mechanism (ERM) of the EMS. Then the correlation between the pound-dollar and the mark-dollar exchange rates will rise. Secondly, a change in U.S. monetary policy such as the 1979 Volcker experiment also raises that correlation, since both the pound and the mark will change in the same way against the dollar. Hence, we need a model that allows for time-variation in the correlations.

Building a multivariate model with time-varying correlations, however, is rather difficult (see Bollerslev, Chou and Kroner (1992)). The reason is that one has to model not only conditional variances, but also all conditional covariances. This can easily lead to an enormous number of parameters.

We propose a new multivariate GARCH model that allows for time-variation in correlations, but which is nevertheless easy to estimate. We first transform the exchange rate changes into their (uncorrelated) principal components and specify a univariate GARCH model for each component. Then we take the inverse of the principal components construction to transform the conditional component moments back in terms of the exchange rate changes themselves. Since this step requires no further estimation, our indirect approach makes multivariate GARCH estimation quite easy. One only has to estimate several univariate GARCH models in the first step.

In a stylized example we show that our multivariate GARCH model can explain the increments in correlation due to the two policy shifts discussed above. We also show empirically that our model improves on existing alternatives, such as the Bollerslev (1990) constant correlations model.

Hence, we use our model to study how exchange rate correlations have changed over time. We find that the correlations between eight main U.S. dollar exchange rates decreased in the first half of the seventies, possibly due to the rather autonomous monetary and fiscal responses of governments to the 1974-1975 period of stagflation (see Krugman and Obstfeld (1991)). When the dollar depreciated strongly in the second half of the seventies, Germany and Japan intervened heavily in the foreign exchange market in 1977-1978. Together with the inception of the EMS in 1979, this marks a period of greater international policy coordination, and our results show that exchange rate correlations indeed increased.

The correlations remained high during the eighties. This is mainly caused by the huge swing in the dollar against all other currencies under consideration and the EMS stability.

The nineties are characterized by lower correlations, partly due to the EMS crash in 1992. However, in the second half of the nineties there is an upward tendency in the correlation of the mark-dollar rate with the lira-dollar, but not with the pound-dollar rate. This may be due to the advent of EMU, which fixes the mark-lira but not the mark-pound rate.

Chapter 6: Why is it so Difficult to Find an Effect of Exchange Rate Risk on Trade?

Chapter 6 discusses the relation between exchange rates and international (goods) trade. More specifically, we focus on the effect of exchange rate risk on trade. This issue has important implications for several policy discussions. For example, it is a determinant for the choice of an international monetary system, in particular the choice between a fixed and floating exchange rate system. In that respect, it is one of the main economic arguments in favor of monetary unification in Europe, as it is commonly believed that exchange risk has a negative effect on trade.

Also within a floating regime the effect of exchange risk on trade is important. For example, it provides a rationale for foreign exchange interventions, such as those following the 1987 Louvre Accord. After all, one of the motives for intervention is to reduce exchange rate fluctuations, because exchange rate risk is assumed to have an adverse effect on trade (see Edison (1993) and Almekinders and Eijffinger (1991)).

Hence, not surprisingly, the effect of risk on trade has attracted much researchers in international economics. The voluminous theoretical and empirical literature on this topic, however, has produced ambiguous findings (see Côté (1994)).

In Chapter 6 we try to explain why the empirical literature has not given a conclusive answer. We first re-examine the effect of risk on trade for our data set, which concerns bilateral aggregate U.S. exports to the other G7 countries. We try to improve on existing studies by proposing a more accurate real exchange rate risk measure and by paying special attention to the lag structure of the trade equation. Our results on the effect of exchange risk on trade confirm the ambiguity found in the literature.

Next, we analyze why it is so difficult to find a clear effect. The estimates show that export decisions are mostly affected by the exchange rate distribution about one year later. The riskiness of the exchange rate at such a long horizon appears fairly constant over time with only short-term fluctuations. This makes it so difficult to discover the true effect of risk on trade from the limited time series data that are typically available.

It is clear from this overview that Chapters 2 to 6 yield a number of new results, both regarding economic questions and econometric modeling. In Chapter 7 we briefly summarize these results, relate them to each other and provide suggestions for future research. A Dutch summary of the thesis can be found at the end of the book.

Chapter 2

Improving GARCH Volatility Forecasts with a New Regime-Switching GARCH Model

This chapter concerns forecasting the volatility of exchange rates. Many researchers use GARCH models to generate volatility forecasts. We show that such forecasts are too variable. This may be due the high estimated persistence of shocks in GARCH forecasts. To obtain more flexibility regarding volatility persistence, we extend the GARCH model by distinguishing two regimes with different volatility levels; GARCH effects are allowed within each regime. Our specification improves on existing regime-switching GARCH models, for instance by making multi-period-ahead volatility forecasting a convenient recursive procedure. The empirical application on U.S. dollar exchange rates shows that our model yields better volatility forecasts than single-regime GARCH.

2.1 Introduction

Volatility of financial returns is an important aspect of many financial decisions. For example, volatility of exchange rates is a determinant for pricing currency options and it may also affect international goods trade. Hence, there is a need for good volatility forecasts.

Volatility forecasts are often based on the fact that volatility is time-varying in high-frequency data and that periods of high volatility tend to cluster. To capture this, many authors use generalized autoregressive conditional heteroskedasticity

(GARCH) models, introduced by Engle (1982) and Bollerslev (1986); see Bollerslev, Chou and Kroner (1992) for an overview of the GARCH literature. Such models usually improve the fit a lot compared with a constant variance model and, as Andersen and Bollerslev (1998) claim, GARCH models provide good volatility forecasts.

In this chapter we show that GARCH forecasts are, nevertheless, too variable. This may be caused by the well-known high estimated persistence of individual shocks in GARCH volatility forecasts. After all, if large shocks increase subsequent volatility forecasts for a long time, then the variability of the forecasts rises. The main goal of this chapter is to improve on the GARCH volatility forecasts by introducing more flexibility regarding volatility persistence. Therefore, our model contains two regimes with different volatility levels. Persistence of both regimes yields an extra source of volatility persistence compared to standard, single-regime GARCH. Moreover, GARCH formulas are used within each regime and these formulas are allowed to be different, so that volatility persistence can be different across regimes. The additional flexibility implies that a shock can be "relieving pressure" on the system, that is, a shock can be followed by a period of low instead of high volatility, because the process can switch to the low-volatility or low-persistence regime.

The way we average out the unobserved volatility regimes represents a new way of combining a regime-switching model with GARCH into a regime-switching GARCH model. The most important advantage over existing variants is that multi-period-ahead volatility forecasting is a convenient first-order recursive procedure. This enables us to compare standard, single-regime GARCH with regime-switching GARCH forecasts. The empirical study on daily U.S. dollar exchange rates versus the British pound, German mark and Japanese yen from January 1978 to July 1997 shows that regime-switching GARCH out-of-sample forecasts do not suffer from excess variability and outperform standard GARCH forecasts, both one-period and multi-period-ahead.

The remaining part of this introduction presents a brief overview of the regime-switching volatility literature and explains the contribution of this chapter in more detail.

The high volatility persistence of shocks in standard GARCH, apparently the reason behind the excess variability of GARCH forecasts, is well-known from the existing GARCH literature. For instance, for their stock return data Hamilton and Susmel (1994) show that a shock this week will have a nonnegligible effect on the variance a full year later. Lamoureux and Lastrapes (1990), among others, show that the high estimated volatility persistence in GARCH models may originate from structural changes in the variance process. They demonstrate that any shift in the unconditional variance is likely to lead to misestimation of the GARCH parameters in such a way that they imply excess volatility persistence. For example, if the variance is high but constant for some time and low but constant otherwise, persistence of such high- and low-volatility homoskedastic periods already results in volatility persistence. A GARCH model, which cannot capture persistence of such periods, puts all volatility persistence in the persistence of individual shocks.

Note that this idea is similar to Perron's (1989) work on the mean equation. He finds that structural breaks in the mean make it more difficult to reject the null of a unit-root (permanent persistence of shocks in the mean) for an actually trend-stationary process, if the test does not account for breaks.

Structural changes in the variance process can originate from changes in economic policy. For example, the inception of the Exchange Rate Mechanism of the European Monetary System (EMS) in 1979 stabilized intra-European exchange rates. Sudden shifts may also result from more exogenous changes in the economic environment, such as the OPEC oil crises.

One possibility to allow for periods with different unconditional variances is, of course, by introducing deterministic shifts into the variance process, but this is rather ad hoc. A popular approach to endogenize changes in the data generating process is the use of a Markov regime-switching model. Hamilton (1989) introduces this model to describe the U.S. business cycle, which is characterized by periodic shifts from recessions to expansions and vice versa. In our context of exchange rate volatility, a Markov process can be used to govern the switches between regimes with different variances. Even a simple regime-switching model with constant regime-specific variances already captures much of the volatility persistence, as our empirical results show.

To capture the remaining conditional heteroskedasticity, the Markov approach can be combined with ARCH models, as Cai (1994) and Hamilton and Susmel (1994) show. Their regime-switching ARCH models have two ways to capture volatility persistence, namely the persistence of regimes and the ARCH effects within a regime. We show empirically that for some series a moderate number of ARCH terms is indeed sufficient to capture all conditional heteroskedasticity.

A potential drawback of regime-switching ARCH models is that only ARCH models are allowed within a regime, not GARCH models. Our empirical results show that this is not only a theoretical disadvantage. It is also important from a practical point of view, as some series are characterized by long persistence of shocks within a regime. One GARCH term can capture such long persistence much more parsimoniously than a large number of ARCH terms. Moreover, as we show, neglecting the long persistence can result in even worse volatility forecasts than those generated by a single-regime GARCH model. For that reason Gray (1996a) modifies the approaches of Cai (1994) and Hamilton and Susmel (1994) such that also GARCH effects are allowed in each variance regime.

Our variance specification also allows for GARCH effects. The crucial difference with Gray's model is the way the unobserved regime indicator is integrated out. We use all available information, whereas Gray uses only part of it. This results in much simpler, first-order recursive variance forecasting formulas. Together with the first-order recursive structure of the likelihood function, this makes our specification very useful from a practical point of view.

In the next section we formally introduce our regime-switching GARCH model. In section 2.3 we describe the data we use in our empirical application and present the empirical results. Section 2.4 concludes.

2.2 Regime-Switching GARCH

In this section we introduce our regime-switching GARCH model, with which we try to improve the standard, one-regime GARCH volatility forecasts. In the first subsection, we describe the model and relate it existing regime-switching (G)ARCH models. In the second one, we present the recursive forecasting formula that makes volatility forecasting in our model quite convenient.

2.2.1 The Model

To describe the model, we need the following notation. Let S_t denote the logarithm of a spot exchange rate at time t, that is, the domestic currency price of one unit of foreign currency. We concentrate on the exchange rate change $s_t = 100(S_t - S_{t-1})$, so that s_t is the percentage depreciation of the domestic currency from time $t-1$ to t.

The regime-switching GARCH model consists of four elements, namely the mean, regime process, variance and distribution. Two of them, the regime process and variance are crucial for interpreting our empirical results, as they are directly related to the difference between our model and standard, one-regime GARCH models.

As the focus of the chapter is on the volatility rather than the mean of exchange rate changes, we assume for simplicity that the conditional mean of s_t is constant:

$$s_t = \mu + \varepsilon_t, \tag{2.1}$$

where the innovation ε_t has zero mean conditional on the information set of the data generating process, which will be defined below.[1] Such a random walk (with drift) specification for exchange rates has become popular since Meese and Rogoff (1983) and MacDonald and Taylor (1992), who stress the empirical quality of the random walk over structural models of exchange rate determination, particularly in the short run.

The regime process represents the main difference between standard GARCH and regime-switching GARCH. As argued in the introduction, the purpose of the regimes with different volatility levels is to capture part of the volatility persistence. This requires that regimes can be persistent, meaning that the regime staying probabilities can be high. We model this as follows. Let r_t be the (unobserved) variance regime at time t, where the first regime is identified as the low-variance one. Let $p_{t-1}(r_t \,|\, \tilde{r}_{t-1}) = p(r_t \,|\, I_{t-1}, \tilde{r}_{t-1})$ denote the probability of going to regime r_t at time t conditional on the information set of the data generating process, which consists of two parts. The first part, I_{t-1}, denotes the information observed by the econometrician, that is $(s_{t-1}, s_{t-2}, \dots)$. The second part, $\tilde{r}_{t-1}$, is the regime path $(r_{t-1}, r_{t-2}, \dots)$, which is not observed by the econometrician. Note that we use the subscript $t-1$ below an operator (probability, expectation or variance) as short-hand notation for conditioning on I_{t-1}.

[1]It is possible to incorporate, for example, autoregressive terms in the conditional mean without making the formulas that follow essentially different.

As in Hamilton (1989), we assume that r_t follows a first-order Markov process with constant staying probabilities

$$p_{t-1}(r_t\,|\,\tilde{r}_{t-1}) = p(r_t\,|\,r_{t-1}) = \begin{cases} p_{11} & \text{if } r_t = r_{t-1} = 1 \\ p_{22} & \text{if } r_t = r_{t-1} = 2. \end{cases} \tag{2.2}$$

If p_{11} and p_{22} are high, this specification results in the regime persistence required above.

The specification of the conditional variance, the third element of the model, represents the main difference between this chapter and earlier papers on regime-switching (G)ARCH. We will now discuss four specifications of the variance of interest, $V_{t-1}\{\varepsilon_t\,|\,\tilde{r}_t\}$; the final one turns out to be the most convenient.

The first specification of the conditional variance is a direct application of the GARCH(1,1) model in a regime-switching context:

$$V_{t-1}\{\varepsilon_t\,|\,\tilde{r}_t\} = \omega_{r_t} + \alpha_{r_t}\varepsilon_{t-1}^2 + \beta_{r_t}V_{t-2}\{\varepsilon_{t-1}\,|\,\tilde{r}_{t-1}\}, \tag{2.3}$$

where the current regime only determines the parameters, that is, the intercept ω_{r_t}, the ARCH parameter α_{r_t} and the GARCH parameter β_{r_t}. This specification, however, appears practically infeasible when estimating the model. This is due to the fact that $V_{t-1}\{\varepsilon_t\,|\,\tilde{r}_t\}$ in (2.3) depends on the entire regime path $\tilde{r}_t$, because it depends on r_t and $V_{t-2}\{\varepsilon_{t-1}\,|\,\tilde{r}_{t-1}\}$, which depends on r_{t-1} and $V_{t-3}\{\varepsilon_{t-2}\,|\,\tilde{r}_{t-2}\}$, which depends on r_{t-2} and $V_{t-4}\{\varepsilon_{t-3}\,|\,\tilde{r}_{t-3}\}$, and so on. Since the number of possible regime paths grows exponentially with t, this leads to an enormous number of paths to t. The econometrician, who does not observe regimes, has to integrate out all possible paths when computing the sample likelihood. This renders estimation intractable. The remaining specifications of the conditional variance are ways to avoid this problem.

The second specification is based on Cai (1994) and Hamilton and Susmel (1994). They essentially remove the GARCH term, which is the cause of the path dependence, and thus use only an ARCH term in (2.3). Since $V_{t-1}\{\varepsilon_t\,|\,\tilde{r}_t\}$ then only depends on the current regime r_t, there is no problem of path dependence.[2]

[2] Cai (1994) and Hamilton and Susmel (1994) use slightly different models in which $V_{t-1}\{\varepsilon_t\,|\,\tilde{r}_t\}$ not only depends on the current but also on a few recent regimes. The essential point is that the conditional variance depends only a small number of regimes, which can be integrated out in the likelihood quite easily.

The third specification of the conditional variance comes from Gray (1996a). He argues that the problem of path dependence can be solved without giving up the potentially important persistence effects of a GARCH term, as has been done in the second specification. The basic idea of Gray (1996a) is to integrate out the unobserved regime path $\tilde{r}_{t-1}$ directly in the source of the path dependence, $V_{t-2}\{\varepsilon_{t-1} \,|\, \tilde{r}_{t-1}\}$ in (2.3), instead of only in the likelihood. This makes $V_{t-1}\{\varepsilon_t \,|\, \tilde{r}_t\}$ only depend on the current regime r_t, not on the path $\tilde{r}_{t-1}$, as should be clear from our explanation of the path dependency problem below (2.3). Since Gray (1996a) uses the information observable at time $t-2$ when integrating out, he actually assumes that

$$V_{t-1}\{\varepsilon_t \,|\, \tilde{r}_t\} = \omega_{r_t} + \alpha_{r_t}\varepsilon_{t-1}^2 + \beta_{r_t} E_{t-2}[V_{t-2}\{\varepsilon_{t-1} | \tilde{r}_{t-1}\}], \tag{2.4}$$

where the expectation on the right-hand-side is across the regime path $\tilde{r}_{t-1}$, conditional on information I_{t-2}. Note that this is equivalent to integrating out only the single regime r_{t-1}, as the lag of (2.4) implies that $V_{t-2}\{\varepsilon_{t-1} | \tilde{r}_{t-1}\}$ is independent of $\tilde{r}_{t-2}$.

The main benefit of specification (2.4) is that there is no path dependence problem any more, although GARCH effects are still allowed. Furthermore, Gray (1996a) shows that the likelihood function can be computed in a convenient recursive way, similar to that in a single-regime GARCH model.

There is, however, one important problem with Gray's method, especially regarding our focus of volatility forecasting: generating multi-period-ahead variance forecasts $V_{t-1}\{s_{t-1+l}\}$ for some lead $l > 1$ turns out to be very complicated. The approach we will use, which is the fourth and final specification we discuss, does not have this problem with forecasting. Nevertheless, our specification also allows for persistence effects through a GARCH term and also leads to a convenient recursive likelihood function.

The improvement of our model with respect to forecasting is due to two crucial differences with Gray's (1996a) model. First, as the expectation in (2.4) shows, Gray integrates out the regime r_{t-1} at time $t-2$. We postpone this till $t-1$, the time at which we want to compute the conditional variance. This allows us to use more observable information when integrating out the previous regime. This extra data embodies information about previous regimes and is thus useful.

The second crucial difference is that, when integrating out the regime r_{t-1}, Gray does not use the information that the variance regime at time t is in the conditioning information of $V_{t-1}\{\varepsilon_t\,|\,\widetilde{r}_t\}$. Particularly if regimes are highly persistent, r_t gives much information about r_{t-1}. In contrast to Gray, we do use this information.

In formula, our regime-switching GARCH(1,1) model is described by

$$V_{t-1}\{\varepsilon_t\,|\,\widetilde{r}_t\} = \omega_{r_t} + \alpha_{r_t}\varepsilon_{t-1}^2 + \beta_{r_t}E_{t-1}[V_{t-2}\{\varepsilon_{t-1}\,|\,\widetilde{r}_{t-1}\}|r_t], \tag{2.5}$$

where the expectation on the right-hand-side is across the regime path $\widetilde{r}_{t-1}$, conditional on information I_{t-1} and r_t. Note that this is equivalent to integrating out only the single regime r_{t-1}, as the lag of (2.5) implies that $V_{t-2}\{\varepsilon_{t-1}|\widetilde{r}_{t-1}\}$ is independent of $\widetilde{r}_{t-2}$.

To ensure positivity of $V_{t-1}\{\varepsilon_t \mid \widetilde{r}_t\}$ for all t, we assume $\omega_{r_t} > 0$ and α_{r_t}, $\beta_{r_t} \geq 0$. We also assume that the "unconditional" variance $V\{\varepsilon_t \mid r_t\}$ exists for both regimes for all ω_1 and ω_2. Necessary conditions for this are in appendix 2.A, which also provides a formula for the unconditional variance. The necessary conditions are similar to the necessary (and sufficient) condition $\alpha+\beta<1$ in the one-regime GARCH(1,1) model.

By construction, $V_{t-1}\{\varepsilon_t \mid \widetilde{r}_t\}$ in (2.5) only depends on the current variance regime r_t, so that $V_{t-1}\{\varepsilon_t\,|\,\widetilde{r}_t\} = V_{t-1}\{\varepsilon_t\,|\,r_t\}$. Hence, as for Gray's (1996a) specification, there is no path dependence. Moreover, the likelihood function with our specification can also be computed recursively, similar to that in a single-regime GARCH model (see appendix 2.C). In contrast with the Gray (1996a) specification, however, our specification leads to relatively simple, recursive formulas for the conditional variance of future exchange rate changes (see next subsection).

It is clear from above that, even within a regime, we allow for volatility persistence. The exact nature of this persistence is allowed to differ between regimes. For instance, if shocks are more persistent in periods of high volatility than in periods of low volatility, this can be captured by the regime specific parameters in (2.5). This has important consequences for capturing the "pressure relieving" effect of some large shocks. Any regime-switching model can capture this to some extent by a shift from the high-volatility to the low-volatility regime. However, our regime-switching model with different parameters across regimes has a second source of neglecting large recent shocks. If the low-variance regime is the short persistence regime, the large shock will be out of the market very soon after the switch to the low-variance regime. This extra flexibility regarding the volatility persistence of shocks will prove important in the empirical analysis. In this respect, our model generalizes the models in Hamilton and Susmel (1994) and Cai (1994), even if GARCH terms are not present. After all, their regime variances only differ by a multiplicative or additive constant, respectively, not by differences in the ARCH parameters.

The fourth and final element of the regime-switching GARCH model is the conditional distribution. We assume that, conditional on I_{t-1} and $\tilde{r}_t$, the innovation ε_t has a t-distribution with ν_{r_t} degrees of freedom, zero mean and variance $V_{t-1}\{\varepsilon_t | r_t\}$:

$$\varepsilon_t | I_{t-1}, \tilde{r}_t \quad \sim \quad t\,(\nu_{r_t}\,,\, 0\,,\, V_{t-1}\{\varepsilon_t | r_t\})\,. \tag{2.6}$$

The use of a t-distribution instead of a normal one is quite popular in the standard, single-regime GARCH literature (see Bollerslev, Chou and Kroner (1992)). We, however, allow for leptokurtosis even within the regime-specific conditional distribution. We will show empirically that this is important in regime-switching GARCH models, since it improves the stability of the variance regimes. After all, in case of normality, a large innovation in the low-volatility period will lead to a switch to the high-volatility regime earlier, even if it is a single outlier in an otherwise quiet period. Note that the t-distribution includes the normal distribution as the limiting case where the degrees of freedom go to infinity.

Equations (2.1), (2.2), (2.5) and (2.6) describe the complete regime-switching GARCH model. It contains the standard, one-regime GARCH(1,1) model as a special case, since that model results when all regime-specific parameters are equal. The regime-switching GARCH model can be estimated by maximum likelihood (ML). The likelihood function, which has a convenient recursive structure, is derived in appendix 2.C.

2.2.2 Volatility Forecasting

We have claimed before that one of the advantages of our regime-switching GARCH model over Gray's (1996a) model is the ease of multi-period-ahead forecasting. In this subsection we show that forecasting with our model is indeed convenient.

Suppose we need the variance of the l-period-ahead exchange rate change s_{t-1+l} given information available at time $t-1$. For notational convenience, let τ

denote the future time, that is, $\tau = t-1+l \geq t$. The variance of interest is

$$V_{t-1}\{s_\tau\} = \sum_{r_\tau=1,2} p_{t-1}(r_\tau) \cdot V_{t-1}\{\varepsilon_\tau \,|\, r_\tau\}, \tag{2.7}$$

where $p_{t-1}(r_\tau)$ is the probability that the regime at time τ is r_τ conditional on I_{t-1}.[3]

An important implication of our way of modeling the conditional variance in (2.5) is that $V_{t-1}\{\varepsilon_\tau \,|\, r_\tau\}$ in (2.7) can be computed in a first-order recursive manner using a formula analogous to the one Engle and Bollerslev (1986) derived for the standard, one-regime GARCH model. Starting from $V_{t-1}\{\varepsilon_t \,|\, r_t\}$, one can compute $V_{t-1}\{\varepsilon_\tau \,|\, r_\tau\}$ for $\tau > t$ by iterating forward on

$$V_{t-1}\{\varepsilon_{t+i} \,|\, r_{t+i}\} = \omega_{r_{t+i}} + (\alpha_{r_{t+i}} + \beta_{r_{t+i}}) \cdot E_{t-1}[V_{t-1}\{\varepsilon_{t+i-1} \,|\, r_{t+i-1}\} | r_{t+i}] \tag{2.8}$$

for $i = 1, ..., \tau - t$.[4] This makes the computation of $V_{t-1}\{s_\tau\}$ in (2.7) quite easy. The multi-period-ahead volatility forecasts will be compared to the standard, one-regime GARCH(1,1) forecasts in the empirical application of the next section.

2.3 Empirical Results

In this section we empirically examine the quality of the regime-switching GARCH model developed in section 2.2. First, we describe the data. Then we estimate the model and analyze the differences between regime-switching GARCH and standard, single-regime GARCH(1,1). We also compare our regime-switching GARCH model with regime-switching ARCH models similar to the ones used by Cai (1994) and Hamilton and Susmel (1994), so that we can examine whether the GARCH effects, which they neglect, are practically relevant. Finally, we examine whether the introduction of regime-switches to the GARCH model has indeed resulted in better volatility forecasts.

[3] Note that we use the same symbol p_{t-1} for several probabilities (for instance, see (2.2) and (2.7)). The specific meaning of p_{t-1} is uniquely determined by the symbols we use in its argument. This results in a concise notation, which will prove useful in the remaining part of the chapter.

[4] The recursive formula is proved in appendix 2.B, which also provides expressions for the regime probabilities involved in (2.7) and (2.8).

2.3.1 Data

We use three major U.S. dollar exchange rates, namely, the dollar vis-à-vis the British pound, the German mark and the Japanese yen. We have 4,982 daily observations for the exchange rate change s_t from January 3, 1978 to July 23, 1997. All rates have been obtained from Datastream.

Panel A of figures 2.1, 2.2 and 2.3 gives an indication of the volatility clustering of the three exchange rates under consideration over the sample period. We present the squared changes, s_t^2, instead of the changes themselves to get a clearer distinction between periods of high and low volatility. This is also useful when assessing the quality of the volatility forecasts, which will be done below. All three plots show substantial volatility clustering, indicating the usefulness of allowing for conditional heteroskedasticity.

The plots also demonstrate that shocks sometimes have a long effect on subsequent volatility, but that shocks can also be followed by a period of low volatility. To show whether single-regime GARCH, regime-switching ARCH and regime-switching GARCH models can capture this, let us consider figure 2.1A as an example. The large peak in the squared change plot for the British pound on March 27, 1985 was followed by about half a year of substantial volatility. Single and multi-regime ARCH models cannot capture such long-run persistence of individual shocks. However, GARCH models can, and one typically finds a high sum of the ARCH and GARCH parameters in the standard, one-regime GARCH models, indicating high volatility persistence of individual shocks (see Hamilton and Susmel (1994)).

Figure 2.1A also shows that shocks sometimes have the effect of "relieving pressure" on the system, so that a shock is followed by a period of low instead of high volatility. For instance, the G-5 Plaza announcement on September 22, 1985 to bring about a dollar depreciation had a sharp effect on the dollar the next day, as the second largest peak in figure 2.1A makes clear. The sharp fall on that day apparently relieved the foreign exchange market from the tensions that had resulted from the sharp dollar appreciation in the years before, as the foreign exchange market was relatively quiet in the remaining part of 1985. This feature cannot be explained by the large persistence of individual shocks that is typically implied by a standard, one-regime GARCH model. It can, however, be captured by regime-switching models, since these allow for a switch from a

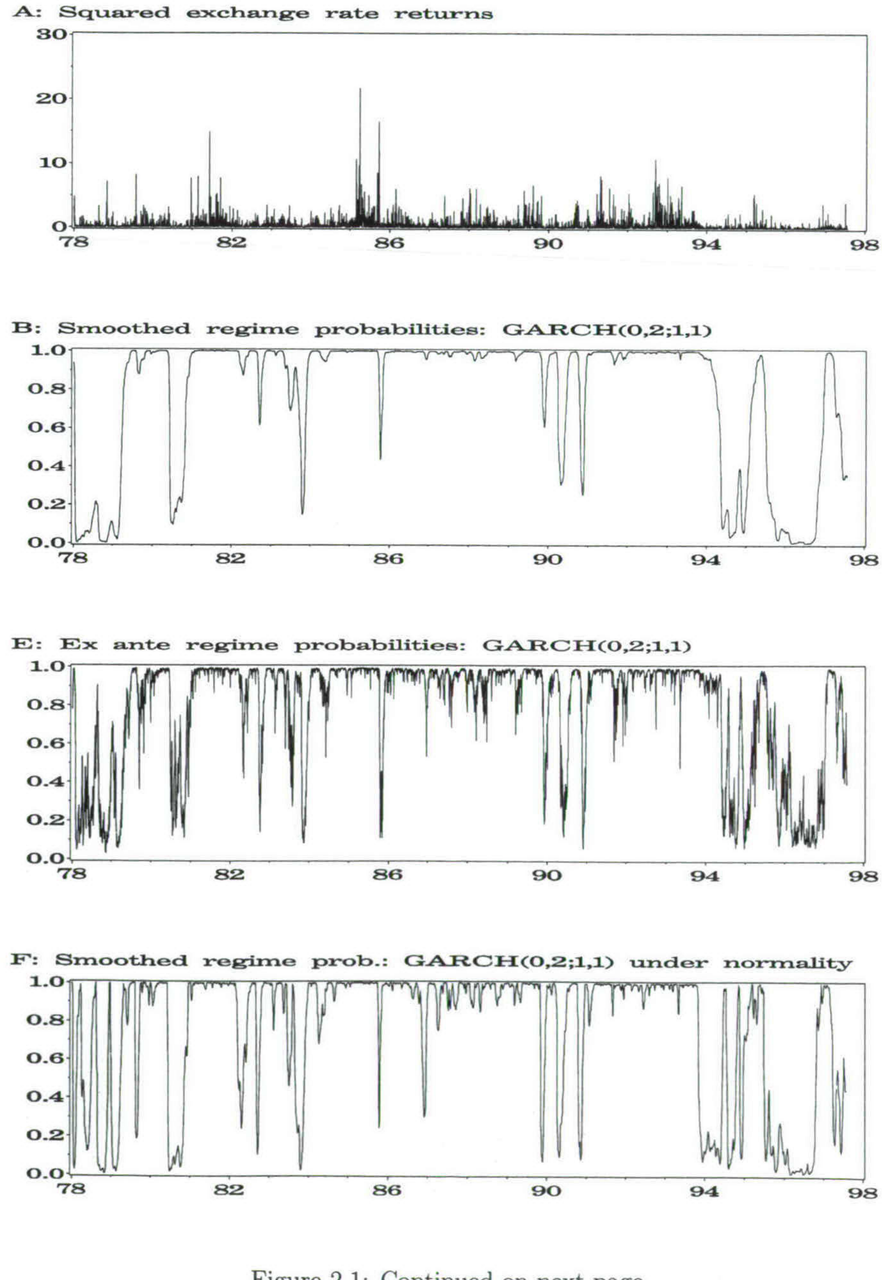


Figure 2.1: Continued on next page

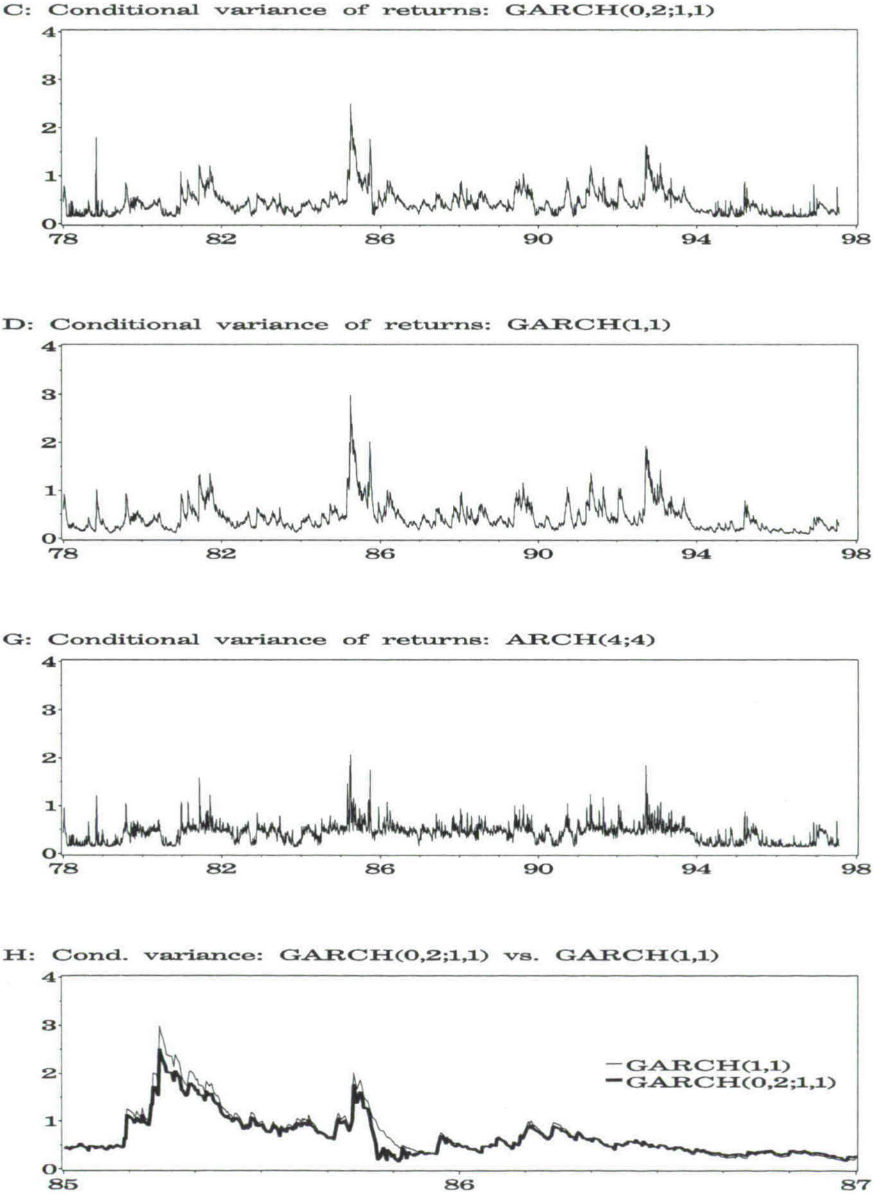


Figure 2.1: British pound over the sample period January 1978 to July 1997

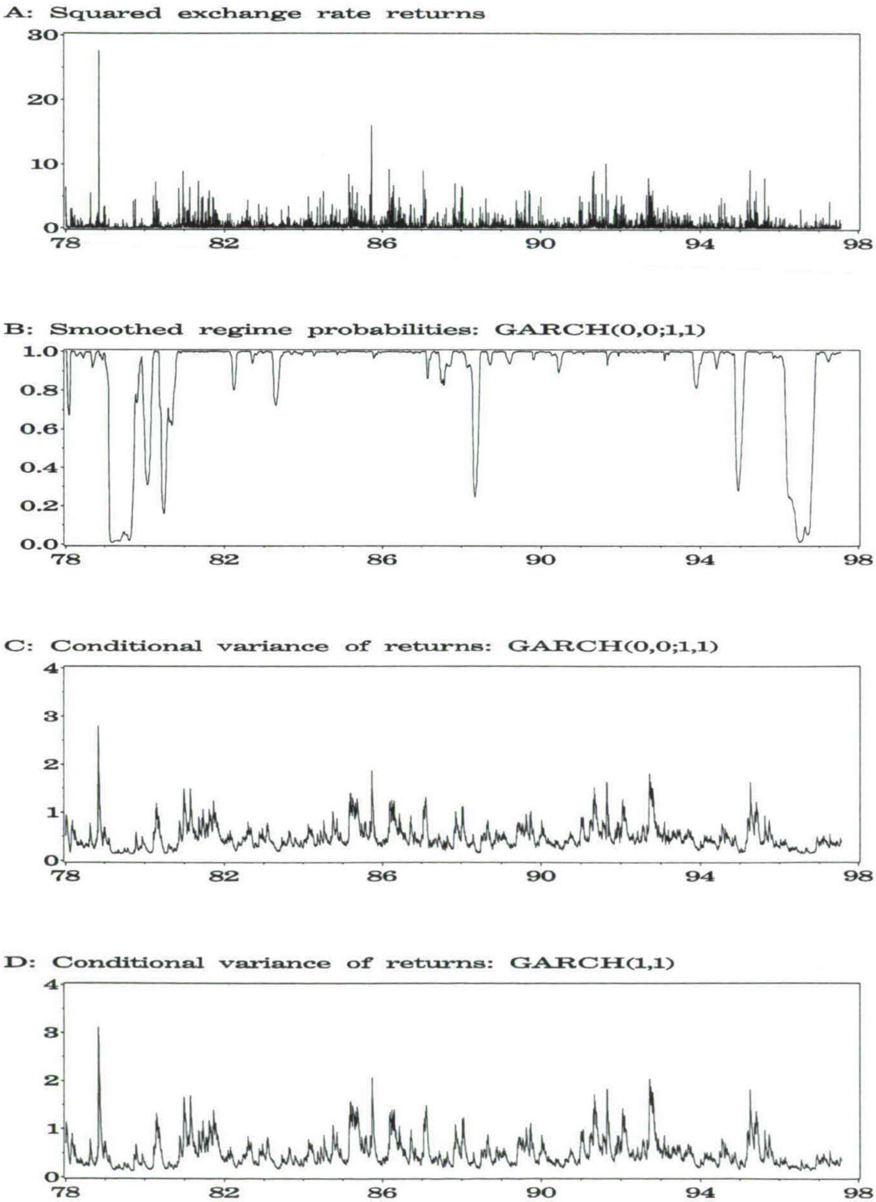


Figure 2.2: German mark over the sample period January 1978 to July 1997

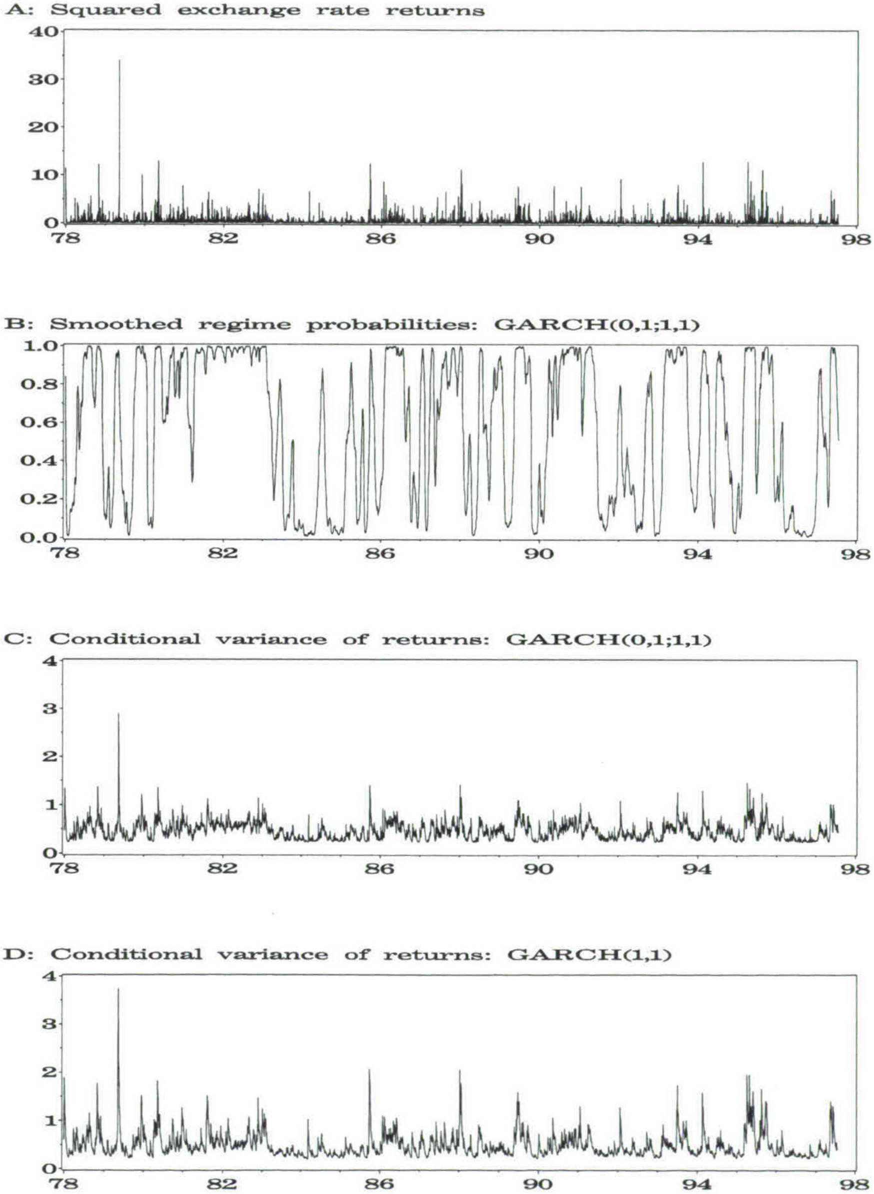


Figure 2.3: Japanese yen over the sample period January 1978 to July 1997

Table 2.1: Moments of exchange rate changes and autocorrelation tests

	British pound	German mark	Japanese yen
Mean	−0.00	0.00	0.01
Variance	0.43	0.47	0.46
Skewness	−0.04	0.04	0.44
Excess Kurtosis	3.04	2.70	3.78
Autocorr. ρ_1	0.07* (0.02)	0.02 (0.02)	0.01 (0.02)
Autocorr. $\widetilde{Q}_{10}$	24.96* [0.01]	6.61 [0.76]	16.16 [0.10]
Autocorr. squares ρ_1^s	0.12* (0.01)	0.12* (0.01)	0.09 * (0.01)
Autocorr. squares Q_{10}^s	539.20* [0.00]	371.24* [0.00]	164.37 * [0.00]

Standard errors in parentheses and p-values in square brackets; * is significant at 5% level.
The first-order autocorrelation, ρ_1, is estimated as the slope coefficient in a regression of the change, s_t, on the first lagged change, s_{t-1}, and a constant. The standard errors are based on White's (1980) heteroskedasticity-consistent asymptotic covariance matrix.
$\widetilde{Q}_{10}$ denotes a modified Box-Pierce type statistic that combines the first ten autocorrelations. Following Pagan and Schwert (1990), it is defined as the sum of the first ten squared normalized autocorrelation estimates, where the normalizing factors are the heteroskedasticity-consistent standard errors of the autocorrelation estimates. $\widetilde{Q}_{10}$ is asymptotically χ^2_{10} distributed.
The first-order autocorrelation in the squared changes, ρ_1^s, and the Box-Pierce type statistic Q_{10}^s are similarly defined, although without the heteroskedasticity correction.

high- to a low-volatility regime in such a situation. Our regime-switching model with GARCH effects can thus capture both the "pressure relieving" effect and the large volatility persistence of shocks, as shown in the previous paragraph.

In table 2.1 we report some descriptive statistics of three exchange rate changes. The second part of the table analyzes the autocorrelation in the changes and their squares. The first-order autocorrelation test for the changes themselves is only significant for the British pound (we always use a significance level of 5%). For that reason, we add a first-order autoregressive term to the mean equation (2.1) for the British pound. Estimates for higher-order autocorrelations are not reported separately, but are combined in Box-Pierce type statistics $\widetilde{Q}_{10}$. Only the one for the pound is significant, but closer inspection of the autocorrelation structure shows that this is mainly due to the first-order autocorrelation. Hence, higher-order autoregressive terms are not necessary.

Table 2.1 also presents two autocorrelation tests for the squared exchange rate changes. Not surprisingly, the tests report clear evidence of conditional heteroskedasticity.

2.3.2 Estimation Results

This subsection presents the estimation results for the regime-switching GARCH model. We use the notation GARCH($P_1, Q_1; P_2, Q_2$) for a regime-switching model with Q_1 (Q_2) ARCH and P_1 (P_2) GARCH terms in the first (second) regime. For comparison, we also estimate four other models. These benchmark models are two single-regime models, namely the constant variance model and the GARCH(1,1) model, and two regime-switching ARCH models, namely the ARCH(0;0) and ARCH(4;4) models (zero and four ARCH terms in both regimes, respectively). The ARCH(4;4) model is comparable with those of Cai (1994) and Hamilton and Susmel (1994). It is, however, somewhat more general, as the ARCH parameters are allowed to differ freely across regimes, whereas in Cai (1994) and Hamilton and Susmel (1994) the parameters only differ by an additive or multiplicative constant, respectively.

Tables 2.2, 2.3 and 2.4 present the maximum likelihood results for the British pound, German mark and Japanese yen, respectively. For easier comparison of the models, we present the "unconditional" variances $\sigma_1^2 = V\{\varepsilon_t \mid r_t = 1\}$ and $\sigma_2^2 = V\{\varepsilon_t \mid r_t = 2\}$ instead of the intercepts ω_1 and ω_2 in the conditional variance formula (2.5). They can be computed from the formulas in appendix 2.A. Moreover, we present the inverse of the degrees of freedom of the t-distribution; testing for conditional normality then boils down to testing whether ν^{-1} differs significantly from zero. Finally, one should be careful when interpreting differences in log-likelihoods in terms of likelihood ratio tests. First, not all models are nested. Second, testing the null of a single-regime against a regime-switching model involves unidentified parameters (the regime-staying probabilities) under the null, so that the likelihood ratio is not asymptotically χ^2 distributed (see Hansen (1992)).

Single-regime GARCH

As is typically found, the standard, one-regime GARCH(1,1) model provides a much better fit than the constant variance model. For instance, the increase in

Table 2.2: Estimation results for the British pound

		SINGLE REGIME		REGIME-SWITCHING		
		Constant variance	GARCH (1,1)	ARCH (0;0)	ARCH (4;4)	GARCH (0,2;1,1)
Mean	μ	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)
Uncond. var. regime 1	σ_1^2	0.46 (0.02)	0.61 (0.18)	0.18 (0.01)	0.19 (0.02)	0.21 (0.05)
ARCH regime 1	α_{11}		0.06* (0.01)		0.13* (0.05)	0.23 * (0.09)
	α_{21}				0.10* (0.04)	0.12 (0.06)
	α_{31}				0.02 (0.03)	
	α_{41}				0.02 (0.04)	
GARCH regime 1	β_1		0.93* (0.01)			
Uncond. var. regime 2	σ_2^2			0.64 (0.03)	0.56 (0.03)	0.52 (0.07)
ARCH regime 2	α_{12}				0.07* (0.02)	0.05 * (0.01)
	α_{22}				0.04 (0.02)	
	α_{32}				0.06* (0.02)	
	α_{42}				0.06* (0.02)	
GARCH regime 2	β_2					0.94 * (0.01)
Degrees of freedom t-dist.	ν_1^{-1}	0.25* (0.02)	0.19* (0.01)	0.20* (0.03)	0.24* (0.03)	0.30 * (0.04)
	ν_2^{-1}			0.15* (0.02)	0.17* (0.02)	0.15 * (0.02)
Regime stay. prob.	p_{11}			0.984 (0.004)	0.992 (0.003)	0.989 (0.005)
	p_{22}			0.986 (0.004)	0.996 (0.002)	0.997 (0.002)
Log-likelihood minus GARCH(1,1)		−4681.60 −244.34	−4437.26 0	−4482.51 −45.25	−4454.37 −17.11	−4412.22 25.04
Autocorr. in squares of	ρ_1^s	0.11* (0.01)	0.01 (0.01)	0.04* (0.01)	−0.01 (0.01)	−0.00 (0.01)
normalized residuals	Q_{10}^s	533.80* [0.00]	5.30 [0.87]	53.92* [0.00]	30.26* [0.00]	9.14 [0.52]

Standard errors in parentheses and p-values in square brackets; * is significant at 5% level.
Both autocorrelation statistics have been defined below table 2.1.
For uniformity with other tables we do not report the (small) first-order autoregressive coefficient that was estimated for the pound only.

Table 2.3: Estimation results for the German mark

		SINGLE REGIME		REGIME-SWITCHING		
		Constant variance	GARCH (1,1)	ARCH (0;0)	ARCH (4;4)	GARCH (0,0;1,1)
Mean	μ	−0.00 (0.01)	−0.01 (0.01)	−0.00 (0.01)	−0.01 (0.01)	−0.00 (0.01)
Uncond. var. regime 1	σ_1^2	0.50 (0.02)	0.70 (0.20)	0.23 (0.02)	0.23 (0.02)	0.14 (0.03)
ARCH regime 1	α_{11}		0.08* (0.01)		0.03 (0.03)	
	α_{21}				0.01 (0.02)	
	α_{31}				0.02 (0.03)	
	α_{41}				0	
GARCH regime 1	β_1		0.91* (0.01)			
Uncond. var. regime 2	σ_2^2			0.76 (0.04)	0.69 (0.05)	0.55 (0.06)
ARCH regime 2	α_{12}				0.06* (0.03)	0.07 * (0.01)
	α_{22}				0.01 (0.02)	
	α_{32}				0.07* (0.03)	
	α_{42}				0.05* (0.02)	
GARCH regime 2	β_2					0.90 * (0.02)
Degrees of freedom t-dist.	ν_1^{-1}	0.25* (0.02)	0.20* (0.01)	0.17* (0.02)	0.19* (0.03)	0.26 * (0.06)
	ν_2^{-1}			0.14* (0.02)	0.15* (0.02)	0.18 * (0.02)
Regime stay. prob.	p_{11}			0.984 (0.004)	0.985 (0.004)	0.981 (0.011)
	p_{22}			0.981 (0.005)	0.987 (0.004)	0.998 (0.002)
Log-likelihood minus GARCH(1,1)		−4962.26 −183.92	−4778.34 0	−4802.59 −24.25	−4791.68 −13.34	−4768.97 9.37
Autocorr. in squares of	ρ_1^s	0.12* (0.01)	0.00 (0.01)	0.03* (0.01)	−0.00 (0.01)	0.00 (0.01)
normalized residuals	Q_{10}^s	371.42* [0.00]	8.03 [0.63]	28.46* [0.00]	27.00* [0.00]	9.41 [0.49]

Standard errors in parentheses and p-values in square brackets; * is significant at 5% level.
A 0 indicates a boundary solution.
Both autocorrelation statistics have been defined below table 2.1.

Table 2.4: Estimation results for the Japanese yen

		SINGLE REGIME		REGIME-SWITCHING		
		Constant variance	GARCH (1,1)	ARCH (0;0)	ARCH (4;4)	GARCH (0,1;1,1)
Mean	μ	−0.01 (0.01)	−0.01 (0.01)	−0.01 (0.01)	−0.01 (0.01)	−0.01 (0.01)
Uncond. var. regime 1	σ_1^2	0.50 (0.03)	0.61 (0.11)	0.23 (0.02)	0.26 (0.03)	0.24 (0.03)
ARCH regime 1	α_{11}		0.09* (0.02)		0.08* (0.04)	0.08 (0.05)
	α_{21}				0	
	α_{31}				0.05 (0.03)	
	α_{41}				0.02 (0.03)	
GARCH regime 1	β_1		0.87* (0.02)			
Uncond. var. regime 2	σ_2^2			0.71 (0.04)	0.68 (0.04)	0.64 (0.05)
ARCH regime 2	α_{12}				0.04 (0.03)	0.06 * (0.02)
	α_{22}				0.05 (0.03)	
	α_{32}				0.05 (0.03)	
	α_{42}				0.01 (0.02)	
GARCH regime 2	β_2					0.78 * (0.10)
Degrees of freedom t-dist.	ν_1^{-1}	0.28* (0.02)	0.25* (0.02)	0.23* (0.03)	0.25* (0.03)	0.26 * (0.03)
	ν_2^{-1}			0.17* (0.02)	0.17* (0.02)	0.18 * (0.02)
Regime stay. prob.	p_{11}			0.976 (0.007)	0.982 (0.006)	0.977 (0.008)
	p_{22}			0.975 (0.006)	0.982 (0.005)	0.983 (0.005)
Log-likelihood minus GARCH(1,1)		−4794.20 −111.62	−4682.58 0	−4672.19 10.39	−4662.58 20.00	−4664.49 18.09
Autocorr. in squares of	ρ_1^s	0.09* (0.01)	0.02 (0.01)	0.04* (0.01)	0.01 (0.01)	0.01 (0.01)
normalized residuals	Q_{10}^s	164.36* [0.00]	13.02 [0.22]	21.57* [0.02]	11.11 [0.35]	11.74 [0.30]

Standard errors in parentheses and p-values in square brackets; * is significant at 5% level.
A 0 indicates a boundary solution.
Both autocorrelation statistics have been defined below table 2.1.

log-likelihood of the GARCH model over the constant variance model is 244.34 for the British pound, so that ARCH and GARCH effects are statistically very important. Furthermore, the sum of the ARCH and GARCH parameters $(\alpha+\beta)$ is large for all three series pointing at high volatility persistence of individual shocks. This has also been found in earlier papers, for instance West and Cho (1995). Panel D of figures 2.1, 2.2 and 2.3 shows the estimated variance series $\widehat{V}_{t-1}\{s_t\}$ for the three series. The volatility persistence appears from the gradual decrease of the conditional variance after a shock.

The high volatility persistence of shocks in the single-regime GARCH model may well indicate parameter instability, as we have shown before. We estimate regime-switching models to analyze whether the high volatility persistence is indeed spurious.

Regime-switching ARCH(0;0)

Let us first consider the regime-switching ARCH(0;0) model in which persistence of regimes is the only source of volatility clustering. Tables 2.2, 2.3 and 2.4 show that all three regime-switching models clearly distinguish between a low- and a high-volatility regime, where the unconditional variance in the latter is about three times as large.

The variance regimes are also persistent, since the staying probabilities p_{11} and p_{22} are all above 0.975. To get a better idea about the amount of persistence that such staying probabilities imply, we compute the expected duration of the high-variance regime. Conditional on being in this regime $(r_t=2)$, this is (see Hamilton (1989))

$$\sum_{l=1}^{\infty} l\cdot P\{r_t{=}2,\ldots,r_{t+l-1}{=}2,r_{t+l}{=}1\,|\,r_t{=}2\} = \sum_{l=1}^{\infty} l\cdot(p_{22})^{l-1}(1{-}p_{22}) = (1-p_{22})^{-1}. \tag{2.9}$$

For a typical ARCH(0;0) staying probability of 0.98, this implies an expected duration of 50 (working) days, which is about 2.5 months.

The introduction of variance regimes captures much of the volatility persistence in the data. To show this, we analyze the normalized residuals. Since, in our model, the variance depends on an unobserved regime, normalizing the residuals is not as easy as usual. One first has to integrate out the regime, as in (2.7). The square root of the resulting variance can then be used as the normalizing factor.

Tables 2.2, 2.3 and 2.4 present tests for heteroskedasticity in the normalized residuals. The first-order autocorrelations ρ_1^s and the aggregate autocorrelation tests Q_{10}^s for the squared normalized residuals show that the conditional heteroskedasticity in the normalized residuals is greatly reduced when going from the constant variance model to the regime-switching model with constant regime-specific variances. However, the heteroskedasticity tests also make clear that there is still heteroskedasticity left. Apparently, there is also volatility clustering within a regime.

Regime-switching ARCH(4;4)

Cai (1994) and Hamilton and Susmel (1994) tried to capture the volatility clustering within regimes by ARCH dynamics. The heteroskedasticity tests for the yen show the usefulness of this approach. A regime-switching ARCH(4;4) model is sufficient to capture all conditional heteroskedasticity in the dollar-yen exchange rate changes, although at the cost of a number of extra parameters.

Allowing for only ARCH effects in a regime-switching model, however, is insufficient for the two European currencies, as the aggregate autocorrelation tests Q_{10}^s show. This remaining conditional heteroskedasticity can be attributed to the high-variance regime, as the ARCH(4;4) estimates for the high-variance regime point at potentially longer persistence in the high-variance regime only; for the low-variance regime they show that even less than four ARCH terms would have been enough. Note that this also illustrates the importance of letting the ARCH parameters differ across regimes. This is in contrast with the models Cai (1994) and Hamilton and Susmel (1994) used, since they restricted the variances in both regimes to be the same apart from an additive of multiplicative scaling parameter, respectively.

An event that appeared to have a particularly long effect on the dollar-pound volatility is the crash of the Exchange Rate Mechanism of the European Monetary System in September, 1992 (see figure 2.1A). Taking only four lags is probably insufficient to generate good conditional variance estimates for this period. The GARCH(1,1) estimated variance in figure 2.1D seems to capture the gradual decrease in volatility better than the ARCH(4;4) variance in figure 2.1G.

Regime-switching GARCH

The long persistence in the high-volatility regime for the pound and the mark, as indicated by the ARCH(4;4) results, show the potential usefulness of GARCH terms as a parsimonious representation of the volatility persistence within regimes. Indeed, the sum of ARCH and GARCH parameters in the high-volatility regime is large for the pound and the mark, and there is no heteroskedasticity left in the normalized residuals. Also, the log-likelihood increases a lot after the introduction of GARCH terms in the high-volatility regime: 42.15 for the pound and 22.71 for the mark. This suggest that the persistence of individual shocks is largest in the dollar-pound exchange rate. For the yen this persistence is smallest, as the log-likelihood increase is -1.91.

Besides the volatility persistence within regimes, the regime-switching GARCH model also uses regime-persistence as a source of volatility persistence. The persistence of regimes can be illustrated by plots of estimated regime probabilities. Following Gray (1996a), we use two types of regime probabilities, namely ex ante and smoothed probabilities. The ex ante probability of a particular regime at time t, is the conditional probability that the process was in that regime at time t using only information available to the econometrician at time $t-1$, that is, I_{t-1}. The smoothed regime probability, on the other hand, uses the complete data set I_T, thereby smoothing the ex ante probabilities.[5] Hence, it gives the most informative answer to the question which regime the process was likely in at time t.

The estimated smoothed regime probabilities of being in the high-variance regime are presented by panel B of figures 2.1, 2.2 and 2.3. We see that the two European currencies have experienced fewer regime shifts than the Japanese yen. Apparently, sudden shifts in the variance are more important for the description of the yen than of the European currencies, where the conditional variance is governed more by smooth transitions (GARCH effects) from high-volatility periods to low ones. This is also clear from the increase in the log-likelihood when introducing regime-specific GARCH effects to the ARCH(0;0) model, which has only

[5]In appendix 2.D we show how to compute the smoothed probabilities in a recursive manner. The algorithm is based on Gray (1996b). It links the ex ante probabilities, which are used during estimation, directly to the smoothed probabilities by iterating forward from the ex ante to the smoothed probabilities.

regime-shifts to capture conditional heteroskedasticity. For the yen the increase is only 7.70, whereas for the pound it is 70.29 and for the mark 33.62.

An issue closely related to the persistence of regimes is the allowance for extra leptokurtosis by a t-distribution within a regime. Without this, the persistence of the, for example, low-volatility regime would have been lower, since then a large sudden change in the exchange rate would have been considered earlier as a shift to the high-volatility regime. This is illustrated by figure 2.1F, which gives the smoothed regime probabilities of the regime-switching GARCH model for the British pound under the restriction of normality. We see that under normality more regime-switches occur.

Comparing single-regime and multi-regime GARCH

So far, we have shown that regime-switching GARCH models are capable of capturing all volatility clustering, whereas regime-switching ARCH models may fail. But standard, one-regime GARCH models also seem to capture the volatility clustering, at least according to the autocorrelation statistics in tables 2.2, 2.3 and 2.4. What is then the reason for introducing an extra source of volatility persistence by allowing for two regimes? Figure 2.1H shows the answer. This plot illustrates the essential difference between single-regime and regime-switching GARCH models. It contains the conditional variance estimates of both GARCH models for the British pound for 1985 and 1986 only. The long effect of the largest shock in the data (March 27, 1985) on subsequent volatility appears to be captured well by both models. However, the sharp fall in the dollar one day after the G-5 Plaza announcement on September 22, 1985 is not dealt with correctly by the one-regime GARCH model. It overestimates the volatility after this event, because of the large volatility persistence of individual shocks. The regime-switching model, however, is able to capture this "pressure relieving" effect, and will thus lead to better volatility forecasts in such periods. It can cut off the effect of large shocks in two ways. First, a switch from the high-volatility to the low-volatility regime leads to a sharp decrease in the variance. This regime-switch is also apparent from figure 2.1E, which plots the ex ante regime probability that the process is in the high-volatility regime (defined above). Second, after the regime-switch shocks have a much shorter effect, as the low-volatility regime is also the low-persistence regime. This second channel is only present if one does not restrict the GARCH parameters in both regimes to be equal. Given the large parameter differences, this channel appears to be an important way to forget large recent shocks.

In summary, our regime-switching GARCH model takes account of two significant aspects of exchange rate distributions, namely the occurrence of many large changes and the clustering of them. The first aspect is already captured in our constant variance model by the t-distribution for the innovations. This characteristic is also present in the four other models. For the second aspect, the clustering of large changes, the regime-switching GARCH model distinguishes two sources. The first one is the persistence of periods with different unconditional variances. This source is absent in the two single-regime models. Secondly, our model allows for volatility persistence even within regimes (in contrast with ARCH(0;0)) and allows for long-run persistence, which is in contrast with ARCH(4;4). In the next subsection we will analyze whether these model differences affect the forecast quality.

2.3.3 Forecasting Performance

In this subsection we compare both the in-sample and out-of-sample forecasts generated by the five models discussed above. The forecasts are computed for two horizons, namely the one-day horizon, which corresponds to the data frequency, and the ten-day horizon.

In-sample forecasting

The in-sample forecasts at time $t-1$ of the variance of the change at some future time τ, $\widehat{V}_{t-1}\{s_\tau\}$, follow from (2.7) after substitution of the estimation results of subsection 2.3.2. Since the conditional variance is the conditional expectation of $(s_\tau-\mu)^2$, we compare $\widehat{V}_{t-1}\{s_\tau\}$ with $(s_\tau-\widehat{\mu})^2$. We first take the one-day-ahead forecasts, so that $\tau=t$.

Following many other papers, such as Gray (1996a), the first forecast statistic we consider is the root mean squared error (RMSE), defined as the square root of $\frac{1}{T}\sum_{t=1}^{T}\left((s_\tau-\widehat{\mu})^2-\widehat{V}_{t-1}\{s_\tau\}\right)^2$. From panel A of tables 2.5, 2.6 and 2.7 we see that the regime-switching GARCH model outperforms the standard, one-regime GARCH model for all series. Apparently, the introduction of regimes into a GARCH model not only leads to a better fit, but also to better forecasts.

Table 2.5: In-sample volatility forecasting statistics for the British pound

Forecast Statistic	SINGLE REGIME Constant variance	SINGLE REGIME GARCH (1,1)	REGIME-SWITCHING ARCH (0;0)	REGIME-SWITCHING ARCH (4;4)	REGIME-SWITCHING GARCH (0,2;1,1)
Panel A: One-day horizon					
RMSE	0.951	0.919	0.931	0.925	0.917
R^2		0.065	0.039	0.049	0.067
$R^2_{\beta=[0,1]'}$		0.061	0.039	0.049	0.066
β_0		0.07* (0.03)	−0.05 (0.04)	−0.03 (0.04)	0.01 (0.04)
β_1	0.93 (0.06)	0.81* (0.08)	1.12 (0.13)	1.05 (0.11)	0.95 (0.10)
Wald for $\beta = [0,1]'$	1.15 [0.28]	5.55 [0.06]	5.27 [0.07]	0.98 [0.61]	0.89 [0.64]
Autocorr. ρ_1	0.11* (0.02)	0.00 (0.02)	0.05* (0.02)	−0.02 (0.02)	−0.01 (0.02)
Autocorr. $\widetilde{Q}_{10}$	157.61* [0.00]	2.89 [0.98]	33.03* [0.00]	14.66 [0.14]	3.00 [0.98]
Panel B: Ten-day horizon					
RMSE	0.956	0.938	0.943	0.944	0.932
$R^2_{\beta=[0,1]'}$		0.038	0.027	0.024	0.048
β_0		0.10* (0.03)	−0.10 (0.06)	−0.09 (0.05)	−0.00 (0.04)
β_1	0.93 (0.06)	0.72* (0.08)	1.23 (0.18)	1.19 (0.16)	0.96 (0.10)
Wald for $\beta = [0,1]'$	1.16 [0.28]	12.56* [0.00]	4.51 [0.10]	5.47* [0.06]	0.98 [0.61]

Standard errors in parentheses and p-values in square brackets; * is significant at 5% level.
The R^2, $R^2_{\beta=[0,1]'}$, β_0, β_1 and the Wald test for $\beta = [0,1]'$ all relate to regression (2.10).
The R^2 under the restriction $\beta = [0,1]'$, denoted by $R^2_{\beta=[0,1]'}$, has been defined by (2.11).
The standard errors for the estimates of β_0 and β_1 in (2.10) and the asymptotic covariance matrix used in the Wald-statistic for $\beta = [0,1]'$ have been corrected for autocorrelation and heteroskedasticity using the Newey and West (1987) asymptotic covariance matrix. Following West and Cho (1995), we have taken Bartlett weights and have used the same data-dependent automatic lag selection rule. This rule has certain asymptotic optimality properties and was introduced by Newey and West (1994).
The two heteroskedasticity-consistent autocorrelation statistics, ρ_1 and $\widetilde{Q}_{10}$ have been defined below table 2.1. They do not appear in panel B, since the unbiasedness of the forecasts no longer implies that the errors η_τ in (2.10) are serially uncorrelated, as the forecast horizon now exceeds the one day period between observations (overlapping sample).

We also see that for the British pound, the series with the highest volatility persistence of shocks, the outperformance of regime-switching GARCH over single-regime GARCH is smallest and that both our model and the single-regime

Table 2.6: In-sample volatility forecasting statistics for the German mark

	SINGLE REGIME		REGIME-SWITCHING		
Forecast Statistic	Constant variance	GARCH (1,1)	ARCH (0;0)	ARCH (4;4)	GARCH (0,0;1,1)
Panel A: One-day horizon					
RMSE	1.023	1.004	1.001	1.000	1.000
R^2		0.044	0.041	0.043	0.045
$R^2_{\beta=[0,1]'}$		0.037	0.041	0.043	0.044
β_0		0.11* (0.03)	−0.06 (0.04)	−0.01 (0.03)	0.06 (0.03)
β_1	0.94 (0.05)	0.72* (0.06)	1.13 (0.11)	1.02 (0.09)	0.85 * (0.06)
Wald for $\beta = [0, 1]'$	1.49 [0.22]	24.64* [0.00]	4.62 [0.10]	0.50 [0.78]	5.24 [0.07]
Autocorr. ρ_1	0.12* (0.03)	0.03 (0.03)	0.05* (0.02)	0.01 (0.02)	0.02 (0.02)
Autocorr. $\widetilde{Q}_{10}$	147.50* [0.00]	8.30 [0.60]	18.07 [0.05]	17.42 [0.07]	8.13 [0.62]
Panel B: Ten-day horizon					
RMSE	1.020	1.019	1.009	1.010	1.010
$R^2_{\beta=[0,1]'}$		0.003	0.020	0.019	0.019
β_0		0.19* (0.05)	−0.04 (0.07)	−0.06 (0.07)	0.09 (0.06)
β_1	0.93 (0.05)	0.52* (0.09)	1.08 (0.15)	1.10 (0.15)	0.76 * (0.12)
Wald for $\beta = [0, 1]'$	1.64 [0.20]	31.79* [0.00]	0.44 [0.80]	1.02 [0.60]	4.91 [0.09]

Standard errors in parentheses and p-values in square brackets; * is significant at 5% level.
For the definitions of the forecast statistics we refer to the notes below table 2.5.

GARCH model outperform the two regime-switching ARCH models. Apparently, in case of strong volatility persistence of individual shocks, taking account of this by GARCH terms is important.

For the series with less volatility persistence of individual shocks, the dominance of our model over single-regime GARCH is larger, especially for the yen. Moreover, the forecasts generated by regime-switching GARCH and ARCH are of almost equal quality. So, using regimes as a source of volatility persistence is particularly important if the persistence of individual shocks is not very large.

The conclusions continue to hold when we use another, often-used forecast

Table 2.7: In-sample volatility forecasting statistics for the Japanese yen

Forecast Statistic	SINGLE REGIME Constant variance	GARCH (1,1)	REGIME-SWITCHING ARCH (0;0)	ARCH (4;4)	GARCH (0,1;1,1)
Panel A: One-day horizon					
RMSE	1.099	1.093	1.087	1.085	1.086
R^2		0.023	0.023	0.027	0.026
$R^2_{\beta=[0,1]'}$		0.014	0.022	0.027	0.026
β_0		0.15* (0.03)	−0.05 (0.05)	−0.01 (0.05)	0.01 (0.04)
β_1	0.91 (0.05)	0.61* (0.06)	1.08 (0.13)	0.99 (0.13)	0.95 (0.09)
Wald for $\beta = [0,1]'$	3.63 [0.06]	49.90* [0.00]	2.00 [0.37]	0.82 [0.66]	0.76 [0.68]
Autocorr. ρ_1	0.09* (0.04)	0.02 (0.04)	0.05 (0.03)	0.01 (0.03)	0.00 (0.03)
Autocorr. $\widetilde{Q}_{10}$	59.44* [0.00]	12.05 [0.28]	11.88 [0.29]	5.84 [0.83]	6.93 [0.73]
Panel B: Ten-day horizon					
RMSE	1.086	1.092	1.081	1.082	1.082
$R^2_{\beta=[0,1]'}$		−0.006	0.009	0.008	0.008
β_0		0.24* (0.07)	−0.03 (0.08)	−0.01 (0.08)	0.02 (0.08)
β_1	0.91* (0.05)	0.40* (0.13)	1.04 (0.19)	0.98 (0.18)	0.92 (0.18)
Wald for $\beta = [0,1]'$	4.14* [0.04]	35.66* [0.00]	0.46 [0.80]	0.67 [0.72]	0.87 [0.65]

Standard errors in parentheses and p-values in square brackets; * is significant at 5% level.
For the definitions of the forecast statistics we refer to the notes below table 2.5.
A negative $R^2_{\beta=[0,1]'}$ results if the variance of the difference between the squared forecast error and the variance forecast, $V\{(s_\tau - \widehat{\mu})^2 - \widehat{V}_{t-1}\{s_\tau\}\}$, is larger than that of the squared forecast errors only. This can happen if, for instance, variance forecasts are very volatile.

statistic, namely the coefficient of determination, R^2, of the regression

$$(s_\tau - \widehat{\mu})^2 = \beta_0 + \beta_1 \widehat{V}_{t-1}\{s_\tau\} + \eta_\tau. \tag{2.10}$$

This regression is comparable with the ones used by Pagan and Schwert (1990). The quality difference between the regime-switching models and the single-regime GARCH model, however, seems smaller now.

Using the R^2 as a forecast statistic has one drawback. It measures the quality of a linear combination, $\beta_0+\beta_1\widehat{V}_{t-1}\{s_\tau\}$, of the forecast, although one is interested in the quality of the forecast itself. Therefore, we prefer the quality measure

$$R^2_{\beta=[0,1]'} = 1 - \frac{V\left\{(s_\tau-\widehat{\mu})^2 - \widehat{V}_{t-1}\{s_\tau\}\right\}}{V\{(s_\tau-\widehat{\mu})^2\}}, \tag{2.11}$$

which can be viewed as the R^2 under the restriction that $\beta_0=0$ and $\beta_1=1$. This forecast statistic is similar to the R^2-type measure used by Gray (1996a).

The tables make clear that this change from R^2 to $R^2_{\beta=[0,1]'}$ has by far the largest effect for the single-regime GARCH model. The reason will become clear below. The seemingly smaller difference between the regime-switching models and the single-regime GARCH model thus appears to be entirely due to an incorrect quality assessment. Using $R^2_{\beta=[0,1]'}$ instead of the potentially misleading R^2 confirms our earlier conclusions based on the RMSE.

At first sight, it may seem that all models yield bad volatility forecasts, as the R^2 statistics are quite low. However, as Andersen and Bollerslev (1998) argue, it is naive to expect a "high" R^2 from regressions such as (2.10). They demonstrate that, even if $(s_\tau-\widehat{\mu})^2$ is a conditionally unbiased estimator of the variance of interest, $V_{t-1}\{s_\tau\}$, it is a very noisy one, which leads to a low R^2. Using better proxies for the latent variable $V_{t-1}\{s_\tau\}$, they show that (single-regime) GARCH models do provide good volatility forecasts. Nevertheless, we find that regime-switching GARCH forecasts are better.

As stated before, the relatively poor forecasting performance of standard, single-regime GARCH models may well be caused by the high volatility persistence of individual shocks that we found in the previous subsection. Indeed, tables 2.5, 2.6 and 2.7 contain evidence for this claim. To explain this, we analyze regression (2.10) again. If the mean and variance forecasts are (conditionally) unbiased, that is, $\widehat{\mu} = E_{t-1}\{s_\tau\}$ and $\widehat{V}_{t-1}\{s_\tau\} = V_{t-1}\{s_\tau\}$, then regression (2.10) implies that $\beta_0=0$ and $\beta_1=1$ and that the error terms η_τ are serially uncorrelated.

The first two implications of the unbiasedness of the forecasts, $\beta_0=0$ and $\beta_1=1$, are tested both individually and simultaneously using OLS estimates for β_0 and β_1. For reasons of uniformity, all tests are corrected for autocorrelation and heteroskedasticity.[6] The individual tests are robust t-tests, whereas

[6] In case of a ten-day forecast horizon, to be discussed below, the errors in (2.10) are no longer uncorrelated, so that the OLS standard errors have to be corrected; see also the notes below table 2.5.

the simultaneous test is a Wald test using a robust estimate of the asymptotic covariance matrix of the OLS estimator for $[\beta_0, \beta_1]'$.

Tables 2.5, 2.6 and 2.7 make clear that the hypotheses are overwhelmingly rejected for the single-regime GARCH model. All estimates for β_0 are significantly above zero, and all but one estimates for β_1 are significantly below one. Apparently, the GARCH variance estimates are too variable. Note that the rejection of $[\beta_0, \beta_1]' = [0, 1]'$ is exactly the reason behind the large difference between the often-used, but potentially misleading unrestricted R^2 and the better forecast statistic $R^2_{\beta=[0,1]'}$ defined in (2.11).

Regarding the estimates for β_0 and β_1, all regime-switching models do much better, since only one coefficient is significantly different from its hypothetical value. Therefore, not surprisingly, the Wald tests for $\beta = [0, 1]'$ clearly prefer the regime-switching models over the single-regime GARCH model.

The crucial difference between the two types of models is that the regime-switching models have regimes as a means to capture volatility persistence, whereas single-regime GARCH models have to put all volatility persistence in the persistence of individual shocks. Given that the excessive variability of the forecasts disappears after the introduction of another way to capture volatility persistence, we conclude that it is the high persistence in single-regime GARCH models that makes the forecasts too variable.

To analyze the last implication of the unbiasedness of the forecasts, uncorrelated error terms η_τ in (2.10), we compute the first-order autocorrelation in the residuals, ρ_1, and the (heteroskedasticity consistent) modified Box-Pierce statistic $\widetilde{Q}_{10}$ introduced in table 2.1. The models that allow most for volatility clustering, standard GARCH, regime-switching ARCH(4;4) and GARCH, indeed show no significant autocorrelation. The highly significant values for ARCH(0;0) for the pound indicate once again that regime-switches alone are sometimes insufficient to capture all predictability of volatility.

Panel B of tables 2.5, 2.6 and 2.7 presents statistics for the ten-day-ahead forecasts ($\tau = t+9$). Most conclusions for the one-week-ahead forecasts also apply here. So, again the regime-switching GARCH model outperforms the single-regime GARCH model, especially for series with moderate persistence of individual shocks. The regime-switching GARCH model also outperforms the regime-switching ARCH models in case of strong volatility persistence within regimes.

Table 2.8: Out-of-sample volatility forecasting statistics for the British pound

	SINGLE REGIME		REGIME-SWITCHING		
Forecast Statistic	Constant variance	GARCH (1,1)	ARCH (0;0)	ARCH (4;4)	GARCH (0,2;1,1)
Panel A: One-day horizon					
RMSE	0.865	0.835	0.842	0.841	0.834
$R^2_{\beta=[0,1]'}$		0.062	0.047	0.049	0.063
β_0		0.05 (0.03)	−0.05 (0.05)	−0.01 (0.05)	0.01 (0.04)
β_1	0.88 (0.07)	0.84* (0.08)	1.09 (0.14)	0.99 (0.13)	0.93 (0.10)
Wald for $\beta = [0,1]'$	2.51 [0.11]	4.39 [0.11]	3.55 [0.17]	0.87 [0.65]	1.70 [0.43]
Autocorr. ρ_1	0.11* (0.02)	−0.01 (0.02)	0.03 (0.02)	−0.02 (0.02)	−0.00 (0.02)
Autocorr. $\widetilde{Q}_{10}$	121.89* [0.00]	5.65 [0.84]	9.02 [0.53]	9.89 [0.45]	3.59 [0.96]
Panel B: Ten-day horizon					
RMSE	0.868	0.849	0.854	0.854	0.848
$R^2_{\beta=[0,1]'}$		0.043	0.031	0.031	0.046
β_0		0.06 (0.04)	−0.12 (0.07)	−0.18* (0.07)	−0.02 (0.05)
β_1	0.89 (0.07)	0.81* (0.09)	1.24 (0.18)	1.36 (0.20)	0.97 (0.11)
Wald for $\beta = [0,1]'$	2.39 [0.12]	4.89 [0.09]	4.34 [0.11]	10.38* [0.01]	1.95 [0.38]

The first half of the sample has been used for estimation and the second half (2,491 days) for forecasting.
Standard errors in parentheses and p-values in square brackets; * is significant at 5% level.
For the definitions of the forecast statistics we refer to the notes below table 2.5.

Overall, our regime-switching GARCH model is thus the preferred model for in-sample forecasting.

Note that the outperformance of the four heteroskedasticity models over the constant variance model is lower than for the one-day forecast horizon, as all $R^2_{\beta=[0,1]'}$ are now closer to zero. Apparently, the longer the forecast horizon, the less valuable is the information in the information set I_{t-1} for forecasting. This is in line with the well-known fact that conditional heteroskedasticity is lower in lower-frequency data.

Table 2.9: Out-of-sample volatility forecasting statistics for the German mark

Forecast Statistic	SINGLE REGIME Constant variance	SINGLE REGIME GARCH (1,1)	REGIME-SWITCHING ARCH (0;0)	REGIME-SWITCHING ARCH (4;4)	REGIME-SWITCHING GARCH (0,0;1,1)
Panel A: One-day horizon					
RMSE	0.923	0.904	0.904	0.904	0.899
$R^2_{\beta=[0,1]'}$		0.040	0.039	0.040	0.049
β_0		0.12* (0.03)	0.03 (0.05)	0.04 (0.05)	0.06 (0.04)
β_1	0.89 (0.06)	0.68* (0.07)	0.88 (0.13)	0.85 (0.13)	0.81 * (0.09)
Wald for $\beta = [0,1]'$	3.04 [0.08]	19.14* [0.00]	1.73 [0.42]	3.22 [0.20]	5.17 [0.08]
Autocorr. ρ_1	0.14* (0.04)	0.03 (0.04)	0.06 (0.04)	0.06 (0.04)	0.02 (0.04)
Autocorr. $\widetilde{Q}_{10}$	99.50* [0.00]	3.49 [0.97]	7.16 [0.71]	8.64 [0.57]	3.39 [0.97]
Panel B: Ten-day horizon					
RMSE	0.923	0.918	0.909	0.910	0.908
$R^2_{\beta=[0,1]'}$		0.017	0.028	0.027	0.033
β_0		0.14* (0.04)	−0.06 (0.08)	−0.09 (0.09)	0.01 (0.06)
β_1	0.89 (0.06)	0.59* (0.09)	1.07 (0.19)	1.09 (0.20)	0.89 (0.12)
Wald for $\beta = [0,1]'$	3.04 [0.08]	25.51* [0.00]	2.21 [0.33]	5.13 [0.08]	4.17 [0.12]

The first half of the sample is used for estimation and the second half (2,491 days) for forecasting.
Standard errors in parentheses and p-values in square brackets; * is significant at 5% level.
For the definitions of the forecast statistics we refer to the notes below table 2.5.

Out-of-sample forecasting

We now turn to the out-of-sample forecasts. We reestimate the five models for the three exchange rate changes using only the first half of the sample. Holding the parameters fixed, we then use the 2,491 observations in the second half (from October 20, 1987 to July 12, 1997) to generate the volatility forecasts $\widehat{V}_{t-1}\{s_\tau\}$.[7] As before, we take the one-day and ten-day horizons.

[7] We also did the reverse, that is, using the second half of the observations for estimation and the first half of the observations for forecasting. The conclusions are similar and are not reported.

Table 2.10: Out-of-sample volatility forecasting statistics for the Japanese yen

	SINGLE REGIME		REGIME-SWITCHING		
Forecast Statistic	Constant variance	GARCH (1,1)	ARCH (0;0)	ARCH (4;4)	GARCH (0,1;1,1)
Panel A: One-day horizon					
RMSE	1.001	0.998	0.987	0.991	0.991
$R^2_{\beta=[0,1]'}$		0.009	0.025	0.019	0.021
β_0		0.16* (0.04)	−0.14* (0.07)	0.06 (0.06)	0.04 (0.07)
β_1	0.87* (0.06)	0.55* (0.07)	1.26 (0.19)	0.80 (0.12)	0.83 (0.15)
Wald for $\beta = [0, 1]'$	4.18* [0.04]	45.86* [0.00]	9.63* [0.00]	3.91 [0.14]	4.32 [0.12]
Autocorr. ρ_1	0.09* (0.03)	−0.02 (0.03)	0.04 (0.03)	−0.00 (0.03)	−0.03 (0.03)
Autocorr. $\widetilde{Q}_{10}$	69.45* [0.00]	8.12 [0.62]	11.43 [0.33]	12.31 [0.27]	11.24 [0.34]
Panel B: Ten-day horizon					
RMSE	1.001	1.010	0.995	0.998	0.998
$R^2_{\beta=[0,1]'}$		0.005	0.011	0.008	0.008
β_0		0.09 (0.09)	−0.34* (0.17)	−0.36 (0.20)	−0.44 (0.23)
β_1	0.87* (0.06)	0.58* (0.16)	1.66 (0.38)	1.61 (0.42)	1.75 (0.47)
Wald for $\beta = [0, 1]'$	4.18* [0.04]	37.56* [0.00]	6.72* [0.03]	9.18* [0.01]	10.73 * [0.00]

The first half of the sample is used for estimation and the second half (2,491 days) for forecasting.
Standard errors in parentheses and p-values in square brackets; * is significant at 5% level.
For the definitions of the forecast statistics we refer to the notes below table 2.5.

Tables 2.8, 2.9 and 2.10 show that the out-of-sample characteristics are similar to the in-sample ones. Single-regime GARCH models generate forecasts that are again too variable, which is not the case for the models that contain two variance regimes as a second way to capture volatility persistence. Secondly, the RMSE and $R^2_{\beta=[0,1]'}$ demonstrate that the regime-switching GARCH model outperforms the standard, one-regime GARCH model, particularly in case of moderate volatility persistence of shocks, and that regime-switching GARCH again outperforms the two other regime-switching models for the British pound, which is the series with the highest persistence of shocks. The regime-switching forecasts for the other two series do not differ much, although the GARCH model is somewhat better for the German mark and ARCH(0;0) is slightly better for the Japanese yen.

2.4 Conclusion

Standard GARCH models often imply a high degree of volatility persistence of individual shocks. We show that this makes GARCH volatility forecasts too variable. To improve the forecasts, we extend the model such that shocks can, but need not be very persistent. This is achieved by the introduction of two regimes with different levels of volatility; a GARCH formula is used to specify the variance within a regime. The empirical application using three U.S. dollar exchange rates shows the importance of regimes in a GARCH model. The volatility forecasts are no longer too variable and they outperform the standard GARCH forecasts substantially.

The regime-switching GARCH model we develop has several interesting properties. First, it provides a new way to average out the unobserved regime indicator in the conditional variance. Compared to the regime-switching GARCH model introduced by Gray (1996a), our specification makes better use of conditioning information and yields a more convenient recursive multi-period-ahead forecasting formula. It also implies a recursive structure for the likelihood function. Given the improvement over the popular single-regime GARCH model, these characteristics make our specification useful from a practical point of view.

The second interesting property of our regime-switching GARCH model is the allowance for GARCH effects within regimes. The empirical results show that the quality of the volatility forecasts can benefit substantially from them, so that regime-switching GARCH is a worthwhile extension of the existing regime-switching ARCH models.

Third, the model allows for time-varying volatility persistence of shocks. As in single-regime GARCH, shock can be persistent. In contrast to standard GARCH, however, shocks can also be "pressure relieving". This can be due to a switch from the high- to the low-volatility regime or due to a switch from the high- to the low-persistence regime. Our empirical results demonstrate that these two causes are correlated, as in periods of high volatility shocks appear more persistent than in periods of low volatility. Note that this result crucially depends on our allowance for different ARCH and GARCH parameters across regimes.

The fourth and final property of our model is that the error distribution is not restricted by normality; a t-distribution can be used as well. We show that this extra source of leptokurtosis is particularly useful in regime-switching models, since it makes the persistence of regimes less sensitive to outliers.

Given the properties just mentioned, the model has a number of other possible applications. For example, the proposed technique of averaging out unobserved regimes to avoid path-dependence of the likelihood function may also be useful in models that combine switches in the mean with a GARCH variance specification (see Chapter 3). Moreover, regime-switching GARCH volatility forecasts can be used to analyze the effect of volatility on stock returns and to price options, for which volatility assessments are crucial. These applications are left for future work.

Appendices

2.A Unconditional Error Variance

In this appendix we derive an expression for the "unconditional" error variance $V\{\varepsilon_t | r_t\}$ and present three necessary conditions for its existence.

Suppose $V\{\varepsilon_t \mid r_t\}$ exists for all $r_t = 1, 2$ and $\omega_1, \omega_2 > 0$. Using the variance definition (2.5), repeated use of the law of iterated expectations yields

$$\begin{aligned}
&V\{\varepsilon_t | r_t\} \\
&= \omega_{r_t} + \alpha_{r_t} E\{\varepsilon_{t-1}^2 | r_t\} + \beta_{r_t} E\{V_{t-2}\{\varepsilon_{t-1} | r_{t-1}\} | r_t\} \\
&= \omega_{r_t} + \alpha_{r_t} E\{E\{\varepsilon_{t-1}^2 | r_{t-1}, r_t\} | r_t\} + \beta_{r_t} E\{E\{V_{t-2}\{\varepsilon_{t-1} | r_{t-1}\} | r_{t-1}, r_t\} | r_t\} \\
&= \omega_{r_t} + \alpha_{r_t} E\{V\{\varepsilon_{t-1} | r_{t-1}\} | r_t\} + \beta_{r_t} E\{V\{\varepsilon_{t-1} | r_{t-1}\} | r_t\} \\
&= \omega_{r_t} + (\alpha_{r_t} + \beta_{r_t}) \cdot E\{V\{\varepsilon_{t-1} | r_{t-1}\} | r_t\},
\end{aligned} \tag{2.12}$$

where the penultimate equality uses that the distribution of the error given the contemporaneous variance regime does not depend on the future variance regime r_t.

Assuming that $\varepsilon_t | r_t$ is unconditionally homoskedastic, stacking $V\{\varepsilon_t | r_t = 1\}$ and $V\{\varepsilon_t | r_t = 2\}$ yields

$$\begin{bmatrix} V\{\varepsilon_t | r_t = 1\} \\ V\{\varepsilon_t | r_t = 2\} \end{bmatrix} = \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} + \begin{bmatrix} A_{1|1} & A_{2|1} \\ A_{1|2} & A_{2|2} \end{bmatrix} \cdot \begin{bmatrix} V\{\varepsilon_t | r_t = 1\} \\ V\{\varepsilon_t | r_t = 2\} \end{bmatrix}, \tag{2.13}$$

where $A_{i|j} = P\{r_{t-1} = i \mid r_t = j\}(\alpha_j + \beta_j)$. Expressions for the probabilities involved are at the end of this appendix.

Let A be the matrix with elements $A_{i|j}$. Since we have assumed that both unconditional variances exist, $I_2 - A$ is invertible. We get the following one-to-one relation between the unconditional variances and the vector of ω_1 and ω_2:

$$\begin{bmatrix} V\{\varepsilon_t | r_t = 1\} \\ V\{\varepsilon_t | r_t = 2\} \end{bmatrix} = (I_2 - A)^{-1} \cdot \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} \tag{2.14}$$

To obtain necessary conditions for the existence of both variances, we use

$A_{2|1} = \alpha_1 + \beta_1 - A_{1|1}$ and $A_{1|2} = \alpha_2 + \beta_2 - A_{2|2}$ to rewrite $(I_2 - A)^{-1}$ as

$$(I_2 - A)^{-1} = \frac{1}{\det(I_2 - A)} \begin{bmatrix} 1 - A_{2|2} & \alpha_1 + \beta_1 - A_{1|1} \\ \alpha_2 + \beta_2 - A_{2|2} & 1 - A_{1|1} \end{bmatrix}. \tag{2.15}$$

Since the variances are strictly positive for all $\omega_1, \omega_2 > 0$, the four elements of $(I_2 - A)^{-1}$ must be nonnegative and $(I_2 - A)^{-1}$ may not have a zero row. Since $\alpha_i + \beta_i \geq A_{i|i}$ for both regimes $i = 1, 2$, this implies that $\det(I_2 - A) > 0$, so that $1 - A_{1|1} \geq 0$ and $1 - A_{2|2} \geq 0$. However, neither $A_{1|1}$ nor $A_{2|2}$ may be unity; otherwise $\alpha_i + \beta_i - A_{i|i} \geq 0$ for both regimes would imply that $\det(I_2 - A) = -(\alpha_1 + \beta_1 - A_{1|1})(\alpha_2 + \beta_2 - A_{2|2}) \leq 0$, which is not the case.

In summary, we have three necessary conditions for the existence of the "unconditional" variances $V\{\varepsilon_t \,|\, r_t\}$, namely $A_{1|1}, A_{2|2} < 1$ and $\det(I_2 - A) > 0$. So, given the definition of $A_{i|i}$, the sum of the regime-specific ARCH and GARCH coefficients must be less than an inverse probability for both regimes; this inverse exceeds one, but not much if regimes are persistent. Moreover, there is some restriction on a combination of the coefficients across regimes. The three necessary conditions show some similarity to the necessary (and sufficient) condition $\alpha + \beta < 1$ in the standard, one-regime GARCH(1,1) model.

To compute the unconditional error variance in (2.13), we need the probability $p(r_{t-1} \,|\, r_t)$ that the previous regime was r_{t-1} given that the current regime is r_t. Using Bayes' rule, we have

$$p(r_{t-1} \,|\, r_t) = \frac{p(r_t \,|\, r_{t-1}) \cdot p(r_{t-1})}{\sum_{r_{t-1}=1,2} p(r_t \,|\, r_{t-1}) \cdot p(r_{t-1})}, \tag{2.16}$$

where $p(r_t \,|\, r_{t-1})$ is constant (see (2.2)) and the theory of Markov processes gives the unconditional probabilities (see Hamilton (1989)):

$$\begin{aligned} p(r_{t-1}=1) &= \frac{1 - p_{22}}{2 - p_{11} - p_{22}} \\ p(r_{t-1}=2) &= \frac{1 - p_{11}}{2 - p_{11} - p_{22}}. \end{aligned} \tag{2.17}$$

2.B Volatility Forecasting

In this appendix we give an expression for $p_{t-1}(r_\tau)$ in the volatility forecasting formula (2.7) and we prove the recursive formula (2.8).

For the future regime probability in (2.7) we have

$$p_{t-1}(r_\tau) = \sum_{r_{t-1}=1,2} p_{t-1}(r_{t-1}) \cdot p_{t-1}(r_\tau \,|\, r_{t-1}), \tag{2.18}$$

where $p_{t-1}(r_{t-1})$ is discussed in (2.29). For the multi-period-ahead probability on the right-hand-side of (2.18), we form the time-constant Markov transition matrix M:

$$M = \begin{bmatrix} p_{11} & 1-p_{22} \\ 1-p_{11} & p_{22} \end{bmatrix}. \tag{2.19}$$

Using the l-th power of M, the theory of Markov processes states that

$$p_{t-1}(r_\tau \,|\, r_{t-1}) = \left(M^l\right)_{r_\tau r_{t-1}}, \tag{2.20}$$

so that (2.18) can be computed.

In the remaining part of this appendix we prove (2.8), that is,

$$V_{t-1}\{\varepsilon_\tau | r_\tau\} = \omega_{r_\tau} + (\alpha_{r_\tau} + \beta_{r_\tau}) E_{t-1}\left\{V_{t-1}\{\varepsilon_{\tau-1} | r_{\tau-1}\} | r_\tau\right\}, \tag{2.21}$$

where, for notational convenience, the index $t+i$ has been substituted by τ. This formula can be proved by repeatedly using the law of iterated expectations. Using definition (2.5), we get

$$\begin{aligned} V_{t-1}\{\varepsilon_\tau | r_\tau\} &= E_{t-1}\left[V_{\tau-1}\{\varepsilon_\tau | r_\tau\} \,|\, r_\tau\right] \\ &= E_{t-1}\left[\omega_{r_\tau} + \alpha_{r_\tau}\varepsilon_{\tau-1}^2 + \beta_{r_\tau} E_{\tau-1}\left\{V_{\tau-2}\{\varepsilon_{\tau-1} | r_{\tau-1}\} | r_\tau\right\} \,|\, r_\tau\right]. \end{aligned} \tag{2.22}$$

For the ARCH part we get

$$\begin{aligned} E_{t-1}[\varepsilon_{\tau-1}^2 \,|\, r_\tau] &= E\{\varepsilon_{\tau-1}^2 | r_\tau, I_{t-1}\} \\ &= E[E\{\varepsilon_{\tau-1}^2 | r_{\tau-1}, r_\tau, I_{t-1}\} \,|\, r_\tau, I_{t-1}] \\ &= E_{t-1}[E_{t-1}\{\varepsilon_{\tau-1}^2 | r_{\tau-1}\} \,|\, r_\tau], \end{aligned} \tag{2.23}$$

where the last equality uses that the error distribution given the contemporaneous variance regime does not depend on the future variance regime.

For the GARCH part in (2.22) we use similar techniques to obtain

$$
\begin{aligned}
&E_{t-1}[E_{\tau-1}(V_{\tau-2}\{\varepsilon_{\tau-1}\,|r_{\tau-1}\}\,|r_\tau)\,|r_\tau] \\
&\qquad = E[E(V\{\varepsilon_{\tau-1}\,|\,r_{\tau-1},I_{\tau-2}\}\,|\,r_\tau,I_{\tau-1})\,|\,r_\tau,I_{t-1}] \\
&\qquad = E(V\{\varepsilon_{\tau-1}\,|\,r_{\tau-1},I_{\tau-2}\}\,|\,r_\tau,I_{t-1}) \\
&\qquad = E[E(V\{\varepsilon_{\tau-1}\,|\,r_{\tau-1},I_{\tau-2}\}\,|\,r_\tau,r_{\tau-1},I_{t-1})\,|\,r_\tau,I_{t-1}] \\
&\qquad = E[V\{\varepsilon_{\tau-1}\,|\,r_{\tau-1},I_{t-1}\}\,|\,r_\tau,I_{t-1}] \\
&\qquad = E_{t-1}[V_{t-1}\{\varepsilon_{\tau-1}\,|\,r_{\tau-1}\}\,|\,r_\tau].
\end{aligned}
\tag{2.24}
$$

The penultimate equality uses that $I_{\tau-2}$ given $r_\tau, r_{\tau-1}, I_{t-1}$ is independent of r_τ, since the Markov structure implies that the distribution of variance regimes $(r_{\tau-2}, r_{\tau-3}, \ldots)$ conditional on $r_{\tau-1}$ and r_τ is independent of r_τ; this makes the changes $(s_{\tau-2}, s_{\tau-3}, \ldots)$ also independent of r_τ once $r_{\tau-1}$ is given.

Substituting the results for the ARCH and GARCH parts in (2.22) gives formula (2.21).

The required probability in (2.21) is

$$
p_{t-1}(r_{\tau-1}\,|\,r_\tau) = \frac{p_{t-1}(r_\tau\,|\,r_{\tau-1})\cdot p_{t-1}(r_{\tau-1})}{p_{t-1}(r_\tau)}. \tag{2.25}
$$

The switching probability follows from (2.2) and the regime probability $p_{t-1}(r_{\tau-1})$ follows in a similar way as $p_{t-1}(r_\tau)$ in (2.18); the denominator is given by (2.18).

2.C Estimation

We estimate the regime-switching GARCH model by maximum likelihood. In this appendix we derive the likelihood function and show that it has a convenient recursive structure.

To obtain the likelihood function, we first need the density of the exchange rate change at time t conditional on only observable information. Let $p_{t-1}(s_t)$ denote this density evaluated at an exchange rate change equal to s_t. It can be split up as

$$
p_{t-1}(s_t) = \sum_{r_t=1,2} p_{t-1}(s_t\,|\,r_t)\cdot p_{t-1}(r_t). \tag{2.26}
$$

We now discuss how to compute both terms on the right-hand-side.

The first term on the right-hand-side, $p_{t-1}(s_t \,|\, r_t)$, denotes the density of the exchange rate change at time t evaluated at the value s_t conditional on I_{t-1} and on the regime having value r_t. This t-density follows from formulas (2.1), (2.5) and (2.6). It is, however, not straightforward how to compute the conditional variance in (2.5), as this requires integrating out the regime path $\tilde{r}_{t-1}$ in $E_{t-1}[V_{t-2}\{\varepsilon_{t-1} \,|\, \tilde{r}_{t-1}\} | r_t]$. Because $V_{t-2}\{\varepsilon_{t-1} \,|\, \tilde{r}_{t-1}\}$ depends only on r_{t-1}, we just need $p_{t-1}(r_{t-1}|r_t)$, the probability that the previous regime was r_{t-1} given that the current regime is r_t and given the information I_{t-1}:

$$p_{t-1}(r_{t-1} \,|\, r_t) = \frac{p_{t-1}(r_{t-1}) \cdot p_{t-1}(r_t \,|\, r_{t-1})}{p_{t-1}(r_t)}, \tag{2.27}$$

where

$$p_{t-1}(r_t) = \sum_{r_{t-1}=1,2} p_{t-1}(r_{t-1}) \cdot p_{t-1}(r_t \,|\, r_{t-1}). \tag{2.28}$$

The constant switching probability $p_{t-1}(r_t \,|\, r_{t-1})$ follows from (2.2).

The remaining term in (2.27) and (2.28) is $p_{t-1}(r_{t-1})$. This probability is crucial, since all regime probabilities in the chapter can be derived from it. Using similar techniques as in Gray (1996a), the following formula shows that this probability has a first-order recursive structure, which simplifies its computation a lot:

$$\begin{aligned} p_{t-1}(r_{t-1}) &= p_{t-2}(r_{t-1} \,|\, s_{t-1}) \\ &= \frac{p_{t-2}(s_{t-1} \,|\, r_{t-1}) \cdot p_{t-2}(r_{t-1})}{p_{t-2}(s_{t-1})} \\ &= \frac{p_{t-2}(s_{t-1} \,|\, r_{t-1}) \cdot \sum_{r_{t-2}=1,2} p_{t-2}(r_{t-2}) \cdot p_{t-2}(r_{t-1} \,|\, r_{t-2})}{p_{t-2}(s_{t-1})}. \end{aligned} \tag{2.29}$$

Hence, the variables to compute $p_{t-1}(r_{t-1})$ are its previous values $p_{t-2}(r_{t-2})$ and the constant $p_{t-2}(r_{t-1} \mid r_{t-2})$ for $r_{t-2} = 1, 2$, and the previous densities $p_{t-2}(s_{t-1} \,|\, r_{t-1})$ and $p_{t-2}(s_{t-1})$. This makes the computation of $p_{t-1}(r_{t-1})$ a first-order recursive process.

The second term on the right-hand-side of (2.26), $p_{t-1}(r_t)$, is the conditional probability that the current regime is r_t. It is given by (2.28).

Having discussed both terms on the right-hand-side of (2.26), we can now compute the density of interest, $p_{t-1}(s_t)$, being a mixture of two t-densities. This density can then be used to build the sample log-likelihood $\sum_{t=1}^{T} \log(p_{t-1}(s_t))$ with which all parameters in the regime-switching GARCH model can be estimated.

From a practical point of view, it is important to realize that the log-likelihood has a first-order recursive structure, similar to that of a standard, one-regime GARCH(1,1) model. After all, for (2.27) and (2.28) one needs the constant $p_{t-1}(r_t \,|\, r_{t-1})$ and the first-order recursive probability $p_{t-1}(r_{t-1})$ (see (2.29)) for all four combinations of (r_t, r_{t-1}); density (2.26) can then be computed from (2.28), the previous change s_{t-1}, (2.27) and the previous variances $V_{t-2}\{\varepsilon_{t-1}|r_{t-1}\}$ for $r_{t-1} = 1, 2$. This first-order recursiveness of $p_{t-1}(s_t)$ makes the calculation of the sample log-likelihood quite fast. To start up the recursive process, we set the required variables equal to their expectation without conditioning on the information set (appendix 2.A describes how to compute the "unconditional" variance $V\{\varepsilon_t \,|\, r_t\}$, which is used to start-up $V_{t-1}\{\varepsilon_t \,|\, r_t\}$ at $t=1$).

2.D Regime Inference

As stated in footnote 5, the smoothed probability that the regime was r_t at time t, $p_T(r_t)$, can be computed recursively. More generally, any ex post regime probability $p_\tau(r_t)$, for a given future time $\tau \in \{t, t+1, \ldots, T\}$, can be computed in a recursive manner, starting from the ex ante probability $p_{t-1}(r_t)$. In this appendix, we verify that claim.

We can write $p_\tau(r_t)$ for both regimes $r_t = 1, 2$ as

$$
\begin{aligned}
p_\tau(r_t) &= p_{\tau-1}(r_t|s_\tau) \\
&= \frac{p_{\tau-1}(s_\tau|r_t) \cdot p_{\tau-1}(r_t)}{\sum_{r_t=1,2} p_{\tau-1}(s_\tau|r_t) \cdot p_{\tau-1}(r_t)},
\end{aligned}
\tag{2.30}
$$

where $p_{\tau-1}(r_t)$ is known from the previous recursion for all $r_t = 1, 2$, if $\tau > t$. If $\tau = t$, it is known from the estimation process, since then it is simply the ex ante probability given by (2.28).

The second ingredient of (2.30) is the density $p_{\tau-1}(s_\tau|r_t)$ for both regime outcomes. For $\tau = t$ it is known from the estimation process (see appendix 2.C), so that the filter probability, $p_t(r_t)$, follows directly from (2.30). Therefore, let us suppose from now on that $\tau > t$.

Computing $p_{\tau-1}(s_\tau \,|\, r_t)$ requires a number of steps. We first write it as

$$
p_{\tau-1}(s_\tau|r_t) = \sum_{r_\tau=1,2} p_{\tau-1}(s_\tau|r_\tau) \cdot p_{\tau-1}(r_\tau|r_t), \tag{2.31}
$$

where we use that the conditional distribution of s_τ given r_τ does not depend on the earlier regime r_t. This formula itself has two ingredients. The first one is the density $p_{\tau-1}(s_\tau|r_\tau)$ for both regime combinations, which is known from the estimation process.

The second term needed in (2.31) is the $(\tau-t)$-period-ahead regime-switching probability $p_{\tau-1}(r_\tau|r_t)$ for all regime outcomes. Once it has been computed, it should be saved, since it will be needed in the next recursive step. Making use of the Markov structure of the regime process, it can be written in terms of $(\tau-1-t)$-period-ahead switching probabilities

$$p_{\tau-1}(r_\tau|r_t) = \sum_{r_{\tau-1}=1,2} p_{\tau-1}(r_\tau|r_{\tau-1}) \cdot p_{\tau-1}(r_{\tau-1}|r_t). \tag{2.32}$$

Again we have two ingredients. First, we need $p_{\tau-1}(r_\tau|r_{\tau-1})$ for all regime combinations. These are constant and follow from (2.2).

The second ingredient of (2.32) is $p_{\tau-1}(r_{\tau-1}|r_t)$ for all regime combinations. We have

$$\begin{aligned} p_{\tau-1}(r_{\tau-1}|r_t) &= p_{\tau-2}(r_{\tau-1}|r_t, s_{\tau-1}) \\ &= \frac{p_{\tau-2}(s_{\tau-1}|r_{\tau-1}) \cdot p_{\tau-2}(r_{\tau-1}|r_t)}{\sum_{r_{\tau-1}=1,2} p_{\tau-2}(s_{\tau-1}|r_{\tau-1}) \cdot p_{\tau-2}(r_{\tau-1}|r_t)}, \end{aligned} \tag{2.33}$$

where we use that the conditional density of $s_{\tau-1}$ is independent of the earlier regime r_t once $r_{\tau-1}$ is given. We have two ingredients. First, the conditional density $p_{\tau-2}(s_{\tau-1}|r_{\tau-1})$ for both regime outcomes. It is known from the estimation process. Secondly, we need the $(\tau-1-t)$-period-ahead switching probability $p_{\tau-2}(r_{\tau-1}|r_t)$ for all regime combinations. This one was saved during the previous recursion, if $\tau > t+1$. If $\tau = t+1$, it equals one.

This completes the algorithm to compute (2.31), which is the second ingredient of (2.30). For each recursion one has to compute (2.33), use the result to compute (2.32) and use this to compute (2.31). Using this in (2.30) yields the ex post probability $p_\tau(r_t)$ from $p_{\tau-1}(r_t)$. Therefore, starting from the ex ante probability $p_{t-1}(r_t)$ one can recursively compute the ex post probability $p_\tau(r_t)$.

Chapter 3

Long Swings in Exchange Rates: Are They Really in the Data?

In this chapter we test the often-used random walk model against the Markov regime-switching model for three main U.S. dollar exchange rates. The latter model allows for breaks in the trend of exchange rates and hence for long swings. Earlier studies conclude that such swings exist, using Wald based tests on quarterly data. We demonstrate that those tests are not very robust in the strongly nonlinear regime-switching model. Instead, we use the more robust likelihood ratio test, for which we simulate the (non-standard) critical values. Remarkably, the evidence from quarterly data disappears. This is caused by the low data frequency, as likelihood ratios on weekly data are significant. Hence, we eventually conclude that long swings are in the data.

3.1 Introduction

Modeling exchange rates has been a main endeavor for economists. Several structural models of exchange rate determination have been developed, but their empirical validity is often questioned, especially in the short-run (see MacDonald and Taylor (1992) for an overview). Hence, many researchers use the random walk, particularly since Meese and Rogoff (1983), who conclude that random walk forecasts outperform those from structural exchange rate models.

The random walk, however, is unsatisfactory from an economic point of view. It ignores any effect of observed changes in economic policy, and, according to the Lucas (1976) critique, such policy shifts may well affect the exchange rate generating process. For instance, regarding monetary policy, Kaminsky (1993) shows theoretically that a change from a contractionary to an expansionary monetary policy increases the exchange rate depreciation, so that monetary policy switches lead to swings in exchange rates. Moreover, the potential relevance of international policy coordination appears from the 1985 Plaza agreement, in which the G-5 countries announced to try to bring about a U.S. dollar depreciation after the sharp dollar appreciation during the five years before; the dollar indeed depreciated strongly from 1985 to 1987. Both examples show that policy shifts can lead to changes in the trend of exchange rates and thus to long swings.

The idea of long swings is further supported by time plots of exchange rates. For instance, see figures 3.1A, 3.2A and 3.3A in section 3.3, which plot the weekly dollar price of one German mark, Japanese yen and U.K. pound, respectively, from April 1974 to July 1997.

The focus of this chapter is whether long swings exist. This issue is relevant for various reasons. First, if swings exist, this may be an indication of the relevance of economic policy for exchange rate determination, despite the empirical rejections of existing structural exchange rate models. Hence, knowing whether long swing exist is important for future research on exchange rate determination.

Another reason to analyze the existence of long swings is that such swings can provide an explanation for peso problems. In other words, the existence of long swings can explain that exchange rate expectations of rational investors appear biased ex post for a long time. After all, if swings exist, rational investors incorporate the possibility of a swing reversal in their expectations, even though the swing reversal may not materialize in the actual exchange rate process for a long time. This leads to a long period of ex post biased expectations (see also Kaminsky (1993)).

In this chapter we formally examine whether long swings exist for the three dollar rates mentioned above. For that, we test the random walk (with drift) against the more general Markov regime-switching model. The latter model, introduced in the seminal paper of Hamilton (1989), explicitly allows for long swings by defining two regimes with different mean exchange rate changes; persistence of such "mean regimes" leads to the long swings.[1] Hence, the regime-switching

[1] Hence, we now use the regime-switching model for a different purpose than in the previous chapter. There, we used it to describe switches between variance regimes, while allowing for only one mean regime. In Chapter 3, as well as in the next chapter, we use the regime-switching model for switches in mean regimes, while allowing for only one variance regime. The similarity

model is commonly used to test for such swings (see Engel and Hamilton (1990), among others).

Earlier papers such as Engel and Hamilton (1990) conclude that exchange rates exhibit long swings. However, the authors are concerned about the reliability of their Wald based tests in the strongly nonlinear regime-switching model. We show that their tests are indeed not very robust. Hence, we use the more robust likelihood ratio test, for which we simulate the (non-standard) critical values. Remarkably, for similar quarterly data as in Engel and Hamilton (1990), the likelihood ratio tests yield no evidence of swings.

We then analyze whether this lack of evidence is due to the low data frequency by examining weekly exchange rates. To correct for the accompanying conditional heteroskedasticity we introduce a way to extend the basic Hamilton (1989) regime-switching model with a generalized autoregressive conditional heteroskedasticity (GARCH) model for the innovation variance (see Bollerslev, Chou and Kroner (1992) for an overview of GARCH). In contrast to the quarterly data, the likelihood ratios on weekly data are significant. Hence, we eventually conclude that long swings are in the data.

The remaining part of this introduction presents a brief overview of the regime-switching literature and explains the contribution of this chapter in more detail.

In the literature so far, regime-switching models have been used in various ways. Hamilton (1989), Lam (1990), Goodwin (1993), Durland and McCurdy (1994), Filardo (1994) and Ghysels (1994) successfully use regimes to capture recessions and expansions in the U.S. business cycle. Garcia and Perron (1996) and Ang and Bekaert (1998) model interest rates as a regime-switching process.

In contrast to these papers, which focus on the mean of a series, regime-switching models can also be useful to describe the variance. Persistence of regimes with different unconditional variances can explain part of the conditional heteroskedasticity we often find in high-frequency data. Cai (1994), Hamilton and Susmel (1994), Gray (1996a) and Chapter 2 of this book use such "variance regimes" to model the variance of changes in interest rates, stock indices, interest rates and exchange rates, respectively.

---

between Chapters 2 and 3 is that in both chapters we compare regime-switching models with single-regime models, that is, standard GARCH(1,1) in the previous chapter and the random walk in the current one.

Most related to our work are Engel and Hamilton (1990), Kaminsky (1993), Engel (1994) and Dewachter (1997), since they also use regimes to capture long swings in exchange rates. Engel and Hamilton (1990) and Engel (1994) have quarterly data of several major exchange rates, Kaminsky (1993) uses monthly data of the dollar-pound rate, while Dewachter (1997) has weekly data of the dollar versus three European currencies. In all four papers, the authors argue that long swings exist. However, Kaminsky (1993) does not formally test the null hypothesis of a random walk against the regime-switching alternative, while Engel and Hamilton (1990) admit that there is some concern with their Wald test statistics, which are also used in Engel (1994) and Dewachter (1997).

These problems originate from identification problems under the null hypothesis of interest, the random walk. Under this null, only one regime governs the exchange rate, so that the regime-switching probabilities are not identified. This makes the asymptotic distribution of the usual tests (likelihood ratio, Wald and Lagrange multiplier) no longer χ^2, as shown by Hansen (1992).

Engel and Hamilton (1990), Engel (1994) and Dewachter (1997) provide an innovative solution to circumvent this problem by taking the slightly more general null that the current regime is independent of the previous one. This hypothesis implies that there are no long swings, as in the random walk. Under the more general null, however, all parameters are identified, and the authors use Wald statistics to test the hypothesis. Gallant (1987), however, argues that Wald statistics are less robust than, for instance, likelihood ratios in nonlinear models such as regime-switching models. This is clearly illustrated by our computations for the weekly dollar-mark exchange rates: the likelihood ratio for the general null is 9, while the Wald test is extremely high, namely 3,866.

Instead of the Wald based procedure, we test the exact null hypothesis of interest (random walk) against the regime-switching model using the likelihood ratio test; see Lam (1990), Ang and Bekaert (1998) and Garcia (1998) for the application of likelihood ratio tests to other series. The non-standard critical values for our specific model are obtained from Monte Carlo simulations. Remarkably, we find no significant evidence of long swings using similar quarterly data as in Engel and Hamilton (1990) and Engel (1994).

This lack of evidence may be caused by the data frequency: even if swings exist and last for some quarters, sampling at the quarterly frequency may result in too few observations per swing to distinguish the swings from a random walk. This is related to the relevance of sampling frequency for unit root testing. For instance, Choi and Chung (1995) show that increasing the sampling frequency can provide significant improvements in the finite sample power of the augmented Dickey-Fuller unit root test. Therefore, we use monthly and weekly data to enhance the power of the long swings test.

Such high-frequency data exhibit conditional heteroskedasticity. Part of this can be captured by the standard Hamilton (1989) regime-switching model, which allows for different variances across regimes. However, such perfect dependence of mean and variance regimes can be problematic, as the regime process may then be exploited to capture the conditional heteroskedasticity instead of the long swings. This could lead to a significant likelihood ratio test for the existence of regimes, even if long swings do not exist. To avoid this, we disconnect the mean and variance specification and introduce a way to incorporate a GARCH variance structure into the basic regime-switching model.

The empirical results for monthly data do not give evidence of long swings. However, for weekly data they are significant and thus suggest that long swings are a systematic part of the exchange rate generating process.

In the next section, we formally describe the regime-switching model and the procedure to test for long swings. In section 3.3 we describe the data and the empirical results. There, we actually test for the existence of long swings. Section 3.4 concludes.

3.2 Model and Test for Long Swings

To test for the existence of long swings, we first need a model that allows for such swings. In this section we develop that model. We then discuss how to test for the existence of long swings.

3.2.1 Regime-Switching Model

The model we use is an extended version of the Engel and Hamilton (1990) regime-switching model, as we explicitly account for the conditional heteroskedasticity that is present in our weekly data. The model consists of four elements, namely the regime process, mean, variance and distribution. We discuss these elements subsequently, and we relate our specification to the one used by Engel and Hamilton (1990).

We need the following notation. As in Chapter 2, let S_t denote the logarithm of the spot exchange rate at time t, that is, the domestic currency price of one unit of foreign currency. We again concentrate on the exchange rate change $s_t = 100(S_t - S_{t-1})$, so that s_t is the percentage depreciation of the domestic currency from time $t-1$ to t.

The regime process we use is the same as in Engel and Hamilton (1990). It is based on two (unobservable) regimes. Let $r_t \in \{1,2\}$ denote the regime at time t.[2] Within this regime, the mean exchange rate change is μ_{r_t}, which we assume to be constant over time. Across regimes, however, the means are allowed to differ, and we identify the first regime as the low mean regime: $\mu_1 \leq \mu_2$. This provides the basis for the swings. After all, being in the first and then in the second regime for a while leads to a period of appreciation followed by depreciation, that is, to swings in the exchange rate. Note, however, that we do not impose this kind of exchange rate behavior; we do allow for $\mu_1 = \mu_2$, so that exchange rates can have a constant mean.

Whether swings are long or not depends on the regime staying probabilities. Let $p_{t-1}(r_t|\tilde{r}_{t-1}) = p(r_t|I_{t-1}, \tilde{r}_{t-1})$ denote the probability of going to regime r_t at time t conditional on the information set of the data generating process, which consists of two parts. The first part, $I_{t-1} = (s_{t-1}, s_{t-2}, \dots)$, denotes the information that is observed by the econometrician; the second part, $\tilde{r}_{t-1}$, is the regime path $(r_{t-1}, r_{t-2}, \dots)$, which is not observed by the econometrician. Note that we use the subscript $t-1$ below an operator (probability, expectation or variance) as short-hand notation for conditioning on I_{t-1}.

As in Engel and Hamilton (1990), we assume that r_t follows a first-order Markov process with constant staying probabilities, so that

$$p_{t-1}(r_t\,|\,\tilde{r}_{t-1}) = p(r_t\,|\,r_{t-1}) = \begin{cases} p_{11} & \text{if } r_t = r_{t-1} = 1 \\ p_{22} & \text{if } r_t = r_{t-1} = 2. \end{cases} \tag{3.1}$$

Hence, if p_{11} and p_{22} are high, regimes are persistent and exchange rate swings are long.

[2] This symbol r_t should not be confused with the r_t of Chapter 2, which denoted the variance regime, not the mean regime it denotes here.

Although we allow for long swings, or "long-run autocorrelation", there may still be short-run dynamics within a mean regime. In the conditional mean specification we take account of this "short-run autocorrelation" by an autoregressive part, as in Hamilton (1989). We use only one autoregressive term, as it is generally believed that the short-run autocorrelation in exchange rates is small (see West and Cho (1995)):

$$s_t = \mu_{r_t} + \theta(s_{t-1} - \mu_{r_{t-1}}) + \varepsilon_t, \tag{3.2}$$

where the conditional expectation of the innovation is $E_{t-1}\{\varepsilon_t | \tilde{r}_t\} = 0$.

Equations (3.1) and (3.2) are the most important parts of the model, as they relate to the long swings directly. For a complete model specification, however, we also need to define the two other elements, namely the conditional variance of ε_t and its distribution. This is the subject of the remaining part of this subsection.

To specify the conditional variance of ε_t, $V_{t-1}\{\varepsilon_t | \tilde{r}_t\}$, Engel and Hamilton (1990) assume that it is constant within and different across the two mean regimes. The first feature, the constancy within a regime, however, is problematic for our weekly data. After all, if mean regimes are persistent (a few years according to Engel and Hamilton), the variance is also constant for a long time. In particular for high-frequency data this is problematic, as it is well-known that there is conditional heteroskedasticity. Moreover, if the model with constant regime-specific variances is estimated with weekly data, it may well be that the regimes are exploited to capture the strong conditional heteroskedasticity instead of the long swings in which we are interested. This may yield a significant test for the existence of two regimes, even if there are no long swings.

The second feature of the Engel and Hamilton (1990) variance specification is that the variance is different across the two regimes. As the authors admit, the perfect dependence between mean and variance can be problematic. For instance, if the appreciation regime is associated with high volatility, a period of unusual volatility can force the process into this appreciation regime, even when the currency is actually depreciating. Moreover, economists are not convinced that there is any relation between the mean and the variance of exchange rates (for instance, see Engle, Ito and Lin (1990)).

To solve both we disconnect the variance from the regime process, which is thus completely focussed at the long swings. For the variance specification we use the popular GARCH approach, for which we propose a way to incorporate it into a regime-switching model.

A direct application of the standard GARCH(1,1) formula in our regime-switching model would define the conditional error variance as

$$V_{t-1}\{\varepsilon_t|\tilde{r}_t\} = \omega + \alpha\varepsilon_{t-1}^2 + \beta V_{t-2}\{\varepsilon_{t-1}|\tilde{r}_{t-1}\}. \tag{3.3}$$

This specification, however, appears practically infeasible when estimating the model. In building the sample log-likelihood, the econometrician first expresses the unobserved previous surprise term ε_{t-1}^2 in terms of the conditioning variables by using $\varepsilon_{t-1}^2 = \{s_{t-1} - [\mu_{r_{t-1}} + \theta(s_{t-2} - \mu_{r_{t-2}})]\}^2$ (see (3.2)). Hence, $V_{t-1}\{\varepsilon_t|\tilde{r}_t\}$ depends on the unobserved regimes r_{t-1} and r_{t-2}. However, it also depends on the lagged variance $V_{t-2}\{\varepsilon_{t-1}|\tilde{r}_{t-1}\}$, which depends on r_{t-2}, r_{t-3} and $V_{t-3}\{\varepsilon_{t-2}|\tilde{r}_{t-2}\}$, where the latter depends on r_{t-3}, r_{t-4} and $V_{t-4}\{\varepsilon_{t-3}|\tilde{r}_{t-3}\}$, and so on. Consequently, the conditional variance in (3.3) depends on the entire sequence of regimes up to time $t-1$. Since the number of possible combinations grows exponentially with t–1, this leads to an enormous number of regime paths to t–1. The econometrician, who does not observe regimes, has to integrate out all possible regime paths. This renders estimation intractable.

To avoid the path-dependency problem, it is interesting to realize that the same problem also hampered the application of regime-switching GARCH models, where the conditional variance depends on the volatility regime the process is in and where the conditional variance within each regime is governed by a GARCH process. For such models, Gray (1996a) and Chapter 2, in which we adjust Gray's model, introduce a way to remove the path-dependence from the likelihood.

We apply the basic idea behind their techniques to solve the problem also in our regime-switching mean model. That is, we directly average out the regimes r_{t-1} and r_{t-2} in the source of the path-dependence, $\varepsilon_{t-1}^2 = \{s_{t-1} - [\mu_{r_{t-1}} + \theta(s_{t-2} - \mu_{r_{t-2}})]\}^2$, instead of only in the likelihood. This removes the regime-dependence of $V_{t-1}\{\varepsilon_t|\tilde{r}_t\}$. As in Chapter 2, we use the observed information I_{t-1} when averaging out the regimes, so that $V_{t-1}\{\varepsilon_t|\tilde{r}_t\}$ becomes equal to $V_{t-1}\{\varepsilon_t\}$:

$$V_{t-1}\{\varepsilon_t|\tilde{r}_t\} = V_{t-1}\{\varepsilon_t\} = \omega + \alpha E_{t-1}\{\varepsilon_{t-1}^2\} + \beta V_{t-2}\{\varepsilon_{t-1}\}, \tag{3.4}$$

with the usual GARCH restrictions $\omega > 0$ and α, $\beta \geq 0$ to ensure $V_{t-1}\{\varepsilon_t\} > 0$ for all t. We also assume that $\alpha + \beta < 1$, so that the unconditional variance is $\sigma^2 = \frac{\omega}{1-\alpha-\beta}$.

Our feasible GARCH specification (3.4) is more restrictive than the direct but infeasible GARCH application (3.3), as we impose $V_{t-1}\{\varepsilon_t|\widetilde{r}_t\} = V_{t-1}\{\varepsilon_t\}$. However, the only purpose of the variance specification is to make the long swing results robust to conditional heteroskedasticity. Subsection 3.3.4 shows that (3.4) is sufficient for that.

The fourth and final element of the regime-switching model concerns the conditional distribution of exchange rate changes. Engel and Hamilton (1990) choose a normal distribution. However, to allow for extra leptokurtosis in our weekly data, we follow other papers by taking a t-distribution (see Bollerslev, Chou and Kroner (1992)). It has ν degrees of freedom, zero mean, and variance $V_{t-1}\{\varepsilon_t\}$:

$$\varepsilon_t \,|\, I_{t-1}, \widetilde{r}_t \quad \sim \quad t\,(\nu\ ,\ 0\ ,\ V_{t-1}\{\varepsilon_t\})\,. \tag{3.5}$$

Equations (3.1), (3.2), (3.4) and (3.5) describe the complete regime-switching model. As in Engel and Hamilton (1990), we estimate it by maximum likelihood. The likelihood function, which has a convenient recursive structure, is derived in appendix 3.A.

3.2.2 Testing Procedure for Long Swings

The central feature of the model presented in the previous subsection is the allowance for two different mean regimes, as these are able to capture long swings. Hence, to test for the existence of long swings, we test for the presence of two regimes. More formally, we test the null hypothesis of an AR(1) model against the regime-switching model of subsection 3.2.1.[3] In this subsection we set out the testing procedure; the actual empirical test follows in subsection 3.3.2.

The null hypothesis is nested in the regime-switching model, as $\mu_1 = \mu_2$ implies that exchange rates follow an AR(1) process. The likelihood ratio (LR) test is a popular and robust statistic for such cases. However, its asymptotic distribution is no longer χ^2, since the regime-switching probabilities are not identified under the null (see Hansen (1992)).

[3] Thus, we do not literally take the random walk (with drift) as the null model. We add one autoregressive (AR) term to the random walk, just as we do for the regime-switching model (see (3.2)). This is to prevent that short-term autocorrelation contributes to a rejection of the null in favor of the long swings. Given the commonly observed low autocorrelation in exchange rate changes, however, the distinction between the pure random walk and the AR(1) model is not large.

To overcome this problem, Garcia (1998) derives an asymptotic distribution of the likelihood ratio for various commonly used regime-switching models, such as the Engel and Hamilton (1990) model. Although Garcia argues that the asymptotic distributions are potentially invalid, he also simulates the distributions and shows that the asymptotic ones provide good approximations for the true distributions.

Because our regime-switching model with GARCH and t-distributed innovations is different from the models Garcia considers, we cannot use his critical values. Since deriving the correct asymptotic distribution for our model is beyond the scope of this chapter, we directly use simulations to generate the critical values.

The simulation procedure is similar to Lam (1990), Ang and Bekaert (1998) and Garcia (1998). We generate 500 data sets, each containing one series of percentage exchange rate changes s_t for $t = 1, \dots , 1000$. The series are generated under the null restriction $\mu_1 = \mu_2$, which we call μ, so that the data generating process is an AR(1)-GARCH(1,1)-t process. The true parameters for all series are the averages of the parameter estimates for the three real data series of section 3.3, namely $\mu = 0.02$, $\theta = 0.06$, $\sigma^2 = 2.54$, $\alpha = 0.10$, $\beta = 0.88$ and $\nu^{-1} = 0.17$ (we will analyze the sensitivity of the critical values to this choice below). All series start from $s_0 = \mu$. For each series we estimate both the single-regime and the regime-switching AR(1)-GARCH(1,1)-t model, that is, (3.1), (3.2), (3.4), (3.5) with and without $\mu_1 = \mu_2$, respectively. To eliminate or at least reduce the problem of obtaining a local maximum of the regime-switching likelihood function, we use ten different starting vectors and take the maximum of the ten log-likelihoods to compute the likelihood ratio for the data set under consideration (this is similar to Garcia's (1998) approach). This procedure yields 500 likelihood ratios. The critical values are the 10%, 5% and 1% quantiles of their empirical distribution.

The second column of table 3.1 reports the three critical values, namely 6.56, 8.22 and 10.81. They are all much larger than the corresponding standard χ^2_1 critical values 2.71, 3.84 and 6.63, which would apply in the case of one restriction but no unidentified parameters.

With one restriction and two unidentified parameters (p_{11} and p_{22}) under the null, it is interesting to compare the simulated critical values with the standard χ^2_3 values. The latter values are 6.25, 7.81 and 11.34, so that they already seem

Table 3.1: Critical values of lik. ratio test for long swings and sensitivity to DGP

	Basic case	Perturbation to data generating process (DGP)				
		no mean $\mu=0$	no AR $\theta=0$	unit var. $\sigma^2=1$	no ARCH $\alpha=\beta=0$	normality $\nu^{-1}=0$
10% Critical value	6.56	6.56	6.57	6.51	6.51	6.67
5% Critical value	8.22	8.22	8.14	8.21	7.99	8.10
1% Critical value	10.81	10.81	10.64	10.81	12.36	11.18

The null hypothesis is the AR(1) model, which is a special case ($\mu_1 = \mu_2$) of the alternative regime-switching model described in subsection 3.2.1.
The critical values are the corresponding quantiles from the distribution of simulated likelihood ratios; subsection 3.2.2 describes the simulation procedure.

to provide a relatively good indication of the true critical values. This finding is supported by similar results in Lam (1990) and Ang and Bekaert (1998).

Next, we compare our critical values in table 3.1 with the non-standard ones in Garcia (1998). First, our critical values are smaller than the 7.05, 8.68 and 12.00 Garcia obtains for the first-order autoregressive ($\theta=0.337$) homoskedastic ($\alpha = \beta = 0$) normal ($\nu^{-1} = 0$) model.[4] This is probably because our model can capture subsequent large exchange rate innovations of the same sign by the GARCH-t structure, whereas in the homoskedastic normal model they seem to be generated by a mean regime, thereby enlarging the difference between the single-regime and regime-switching likelihoods.

A second comparison with Garcia (1998) concerns his heteroskedastic model. That model is the same as in Engel and Hamilton (1990) ($\theta = 0$, $\alpha = \beta = 0$, $\nu^{-1} = 0$ with two regime-specific constant variances σ_1^2 and σ_2^2, say). Garcia's critical values are much larger (11.88, 13.68 and 17.52). This is due to the crucial difference between the variance specifications of both models. In Garcia (1998) the mean and variance are perfectly dependent, so that the regimes are not only used for the long swings, but may also be exploited to capture the occurrence of high and low volatility periods. This may well lead to high likelihood ratios for the existence of different regimes because $\sigma_1^2 = \sigma_2^2$ is rejected, even though there are no long swings ($\mu_1=\mu_2$). Our model disconnects the variance from the regimes, so that the likelihood ratio only concentrates on the long swings, that

[4]Our results for that model (with $\mu=0.02$, $\theta=0.06$, $\sigma^2=2.54$) are 6.88, 8.44 and 11.98, which are in line with Garcia's critical values.

is, on $\mu_1 = \mu_2$ and not on $\sigma_1^2 = \sigma_2^2$. This makes our critical values lower than those for Garcia's heteroskedastic model. Note that this difference is closely related to the difference in critical values that one would have if no unidentified parameters were present. In that case one would use χ_2^2 values in Garcia's setting and the lower χ_1^2 values in our model, because Garcia's model has an extra null restriction ($\sigma_1^2 = \sigma_2^2$) besides the common restriction $\mu_1 = \mu_2$.

So far, the simulated critical values have been based on one specific set of true parameters. This may be a problem if the critical values are sensitive to the true parameters of the data generating process. Extending Lam (1990) and Ang and Bekaert (1998), we examine this sensitivity by computing the critical values also for five other sets of true parameters (five is already very time consuming). The first set has mean zero ($\mu = 0$), the second has no AR term ($\theta = 0$), the third has unit variance ($\sigma^2 = 1$), the fourth has no ARCH ($\alpha = \beta = 0$), and the fifth has a normal instead of a t-distribution ($\nu^{-1} = 0$). For all sets the other parameters are kept at the averages of the real data estimates, as before. Therefore, each set represents a perturbation of the data generating process (DGP) in only one direction. The rest of the simulation procedure is the same as described above, so we estimate the full single-regime and regime-switching AR(1)-GARCH(1,1)-t models for all parameter sets.

The results in table 3.1 show that the critical values are quite robust to the perturbations, in particular the 10% and 5% values. More specifically, changing the mean μ has no effect, as expected. The low sensitivity regarding the autoregressive parameter θ is in line with Garcia's (1998) results for his regime-switching variants. The low sensitivity with respect to the other parameters is likely due to the fact that our GARCH(1,1)-t specification is separate from the long swings on which the test concentrates.

In summary, we conclude that we can safely use the critical values 6.56, 8.22 and 10.81 for the empirical tests for long swings in the next section.

3.3 Empirical Results

In this section we use the regime-switching model developed above to address the central question of this chapter, namely whether long swings really exist. First, we describe the data. In subsection 3.3.2 we test for the swings. After that, we analyze the estimates of the regime-switching model and in subsection 3.3.4 we present some checks on its specification. In the last subsection, we examine whether taking account of the long swings leads to better exchange rate forecasts than those generated by the random walk model.

3.3.1 Data

We use three U.S. dollar exchange rates, namely, the dollar vis-à-vis the German mark, the Japanese yen and the U.K. pound. These exchange rates have been chosen because of their important role on foreign exchange markets and because they behave relatively independently, for instance, compared to several dollar-EMS exchange rates. We have 1,216 weekly observations for the percentage dollar depreciations s_t over the post-Bretton-Woods period from April 2, 1974 to July 22, 1997. They have been obtained from Datastream. In this subsection we provide some characteristics of the data and use them to motivate our model specification empirically.

In panel A of figures 3.1, 3.2 and 3.3 we show the behavior of the three exchange rates over the sample period. The figures contain the exchange rate levels in U.S. dollars, not in logarithms. At first sight, exchange rates indeed seem to be characterized by long swings.

In table 3.2 we report some descriptive statistics of the three exchange rate changes. There is significant first-order autocorrelation in the weekly German mark changes (we always use a significance level of 5%). For this reason, we have extended the Engel and Hamilton (1990) model by a first-order autoregressive term in the mean equation (3.2). Estimates for higher-order autocorrelations are not reported separately, but are combined in Box-Pierce type statistics $\tilde{Q}_{10}$. They show that higher-order autoregressive terms are unnecessary.

Table 3.2 also presents two autocorrelation tests for the squared exchange rate changes. Both tests point at conditional heteroskedasticity for all three series. This is why we have extended the Engel and Hamilton (1990) model with GARCH specification (3.4) for the conditional error variance.

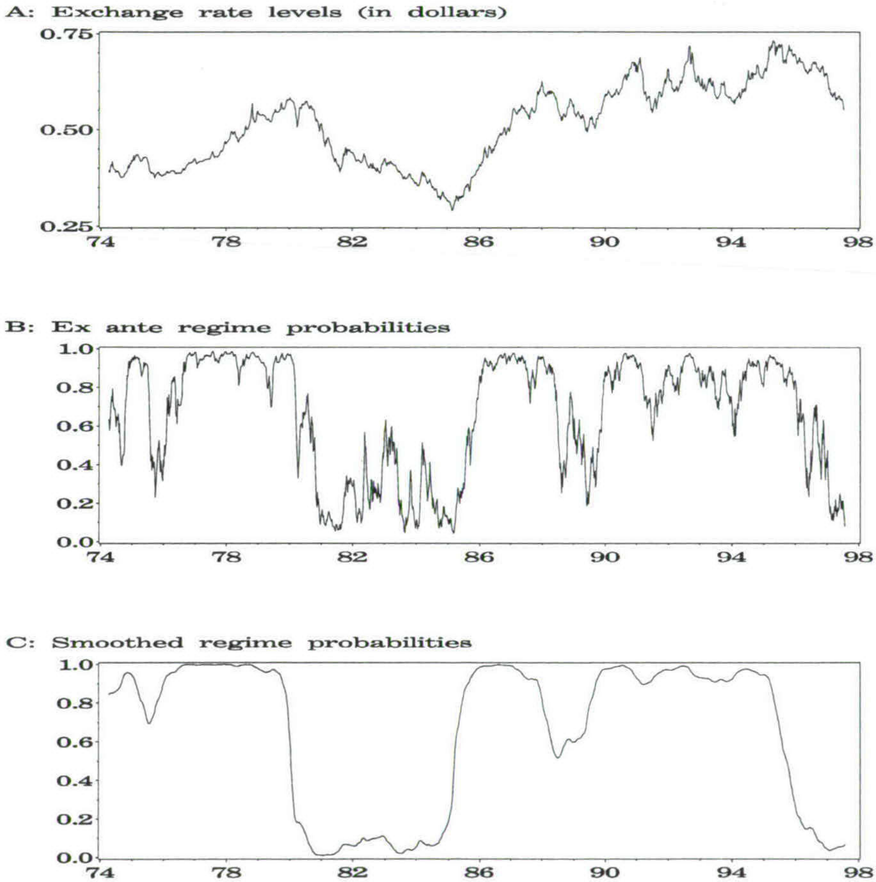


Figure 3.1: German mark over the sample period April 1974 to July 1997

3.3.2 Long Swings in Exchange Rates: Are They Really in the Data?

As we have just seen from figures 3.1A, 3.2A and 3.3A, exchange rates seem to exhibit long swings. In this section we analyze the main point of the chapter, namely whether long swings are a systematic part of the exchange rate generating process, as Engel and Hamilton (1990), Kaminsky (1993), Engel (1994) and Dewachter (1997) claim. After all, the long swings may be only a pattern imposed

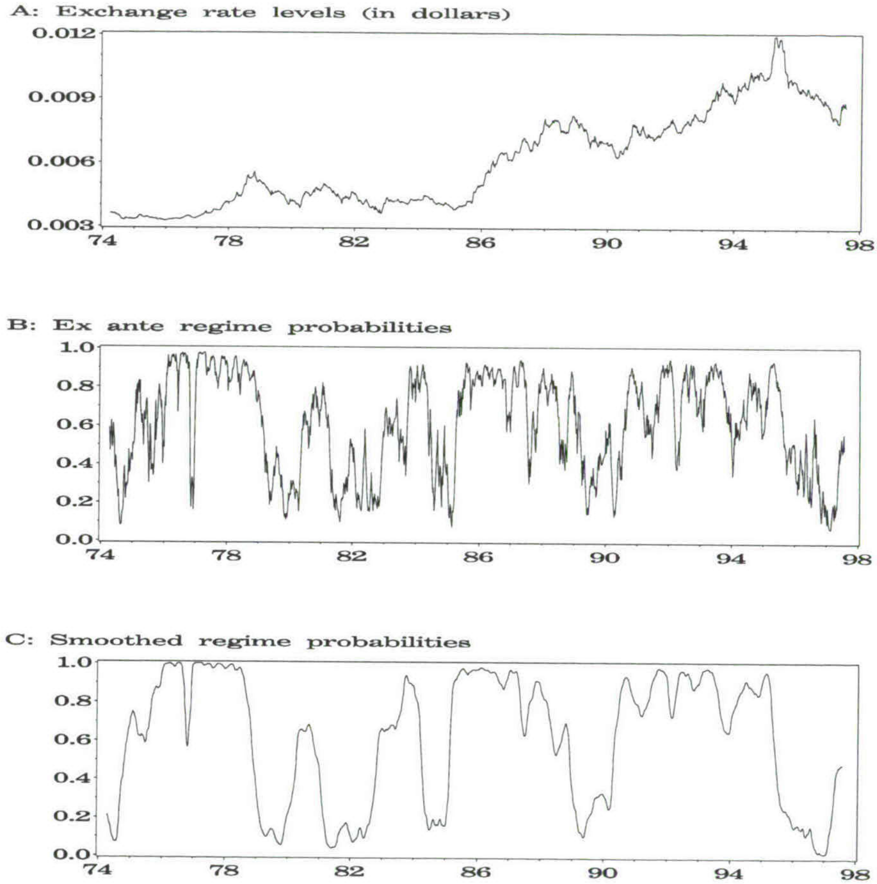


Figure 3.2: Japanese yen over the sample period April 1974 to July 1997

by the eye on the realization of a random walk. As argued in subsection 3.2.2, we take the likelihood ratio test for the existence of two mean regimes in the regime-switching model of section 3.2.1. We first use similar quarterly data as in Engel and Hamilton (1990) and Engel (1994). After that, we enlarge the quarterly series to 1997, and check whether the results change. Finally, we increase the data frequency from quarterly to monthly and then to weekly; this leads to our final answer on the question whether long swings really exist.

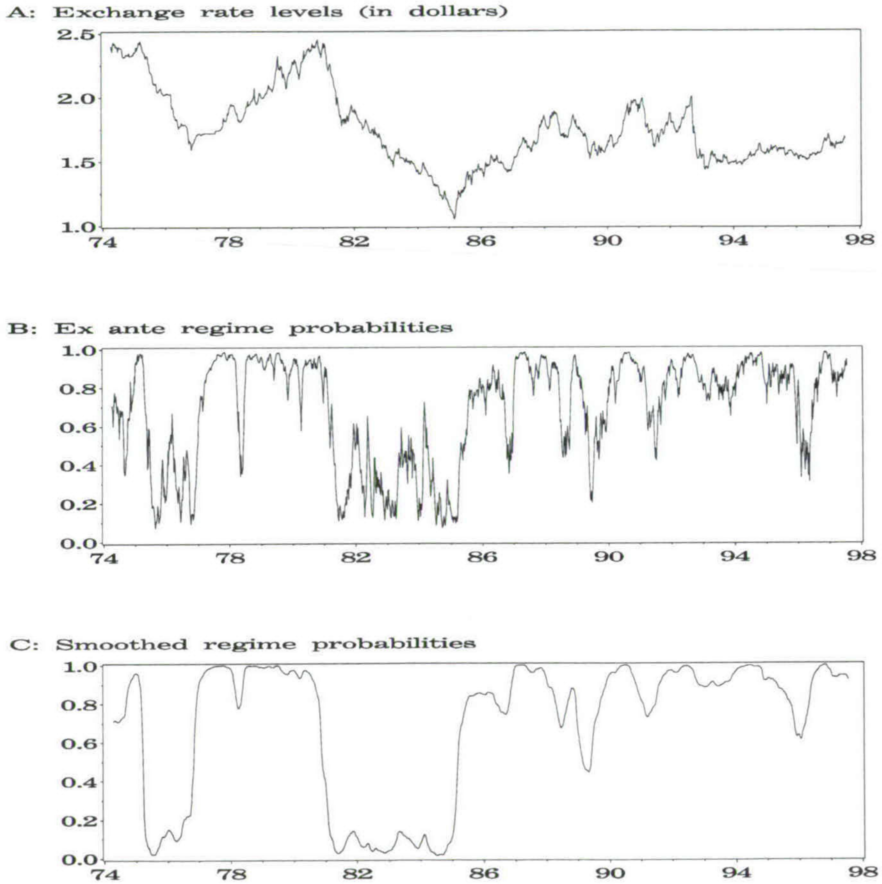


Figure 3.3: U.K. pound over the sample period April 1974 to July 1997

To start our series of tests of the random walk, we take quarterly data from 1974:I to 1987:I, similar to the data used for estimation in Engel and Hamilton (1990) and Engel (1994). The top row of table 3.3 contains the LR tests of the random walk for the three currencies and the 5% critical value. We find significant evidence against the random walk for the U.K. However, the random walk is not rejected for the other two currencies. This latter conclusion is opposite to the one of Engel and Hamilton (1990) and Engel (1994), who claim to have found evidence

Table 3.2: Moments of exchange rate changes and autocorrelation tests

	GERMANY	JAPAN	U.K.
Mean	0.03	0.07	−0.03
Variance	2.14	2.11	2.13
Skewness	−0.14	0.53	−0.40
Excess Kurtosis	1.70	2.01	3.00
Autocorr. ρ_1	0.07* (0.03)	0.05 (0.04)	0.04 (0.04)
Autocorr. $\widetilde{Q}_{10}$	14.07 [0.17]	22.57* [0.01]	6.05 [0.81]
Autocorr. squares ρ_1^s	0.11* (0.03)	0.20* (0.03)	0.20 * (0.03)
Autocorr. squares Q_{10}^s	57.60* [0.00]	92.03* [0.00]	151.82 * [0.00]

Standard errors in parentheses and p-values in square brackets; * is significant at 5% level.
The first-order autocorrelation, ρ_1, is estimated as the slope coefficient in a regression of the change, s_t, on the first lagged change, s_{t-1}, and a constant. The standard errors are based on White's (1980) heteroskedasticity-consistent asymptotic covariance matrix.
$\widetilde{Q}_{10}$ denotes a modified Box-Pierce type statistic that combines the first ten autocorrelations. Following Pagan and Schwert (1990), it is defined as the sum of the first ten squared normalized autocorrelation estimates, where the normalizing factors are the heteroskedasticity-consistent standard errors of the autocorrelation estimates. $\widetilde{Q}_{10}$ is asymptotically χ^2_{10} distributed.
The first-order autocorrelation in the squared changes, ρ_1^s, and the Box-Pierce type statistic Q_{10}^s are similarly defined, although without the correction for heteroskedasticity.

of long swings for the U.K. as well as Germany and Japan.[5] However, Engel and Hamilton admit that there is some concern with their test approach. As discussed in the introduction, they use a Wald test for a slightly more general null than the random walk, so as to circumvent the identification problems associated with the null of a random walk. We have shown in the introduction that the Wald statistic is indeed not the most appropriate statistic to use in a regime-switching model.

So far, we have not found conclusive evidence of the existence of long swings, at least not for Germany and Japan. Of course, this may be due to the absence of swings, but it may also be that a sample period of thirteen years is too short

[5]Although Engel and Hamilton (1990) and Engel (1994) allow for different variances across mean regimes (see subsection 3.2.1), we assume a constant variance over the complete sample period, since we find no conditional heteroskedasticity at the quarterly frequency. Our homoskedasticity assumption is not the reason behind the opposite conclusions; even if we allow for different variances across mean regimes, the likelihood ratio is insignificant.

Table 3.3: Likelihood ratio tests of long swings

Data frequency and period		GERMANY	JAPAN	U.K.	Critical value
Quarterly data	1974-1987	3.17	6.43	8.95*	8.60
	1974-1997	2.42	3.39	5.37	8.60
Monthly data	1974-1997	5.31	4.12	5.37	8.22
Weekly data	1974-1997	9.00*	10.93*	12.16*	8.22

* is significant at 5% level.
The null hypothesis is the AR(1) model, which is a special case ($\mu_1 = \mu_2$) of the alternative regime-switching model described in subsection 3.2.1.
For the quarterly data, we use no autoregressive term ($\theta = 0$) in both the null and alternative model, because there is no autocorrelation at the quarterly frequency. Likewise, we assume conditional homoskedasticity ($\alpha = \beta = 0$). Finally, the error is normally distributed ($\nu^{-1} = 0$), as in Engel and Hamilton (1990) and Engel (1994).
The 5% critical values for the quarterly data are from Garcia (1998); the ones for monthly and weekly data are from the simulation results in table 3.1.

to detect long swings. To analyze this, we extend our sample period by including data from 1987:II to 1997:III. As table 3.3 shows, all LR tests become insignificant now. Hence, our finding based on the quarterly data is that we have no evidence of long swings.

Our inability to reject the random walk can be due to a lack of power of the tests. One reason for this can be that the data frequency is too low. After all, even if swings are part of the exchange rate generating process and last for a number of quarters, quarterly data may result in too few observations per swing to distinguish swings from a random walk. To examine this, we first increase the data frequency from quarterly to monthly. As table 3.3 shows, all LR tests are still insignificant, although they are generally higher than for the quarterly data.

Finally, we use the weekly series described in subsection 3.3.1. The results change completely: all LR statistics are significant now. Hence, the previous inability to find long swings using quarterly or monthly data has statistical reasons: the low data frequency leads to too few observations to significantly distinguish long swings from a single-regime process. Weekly data give the test enough power. Our final conclusion is thus that the data really suggest that long swings exist. Note that this conclusion is entirely based on likelihood ratios, not on the problematic Wald tests that have been used by others.

3.3.3 Estimation Results

We now present the estimation results for our regime-switching model and, for comparison, the results for the random walk. We first consider mean equation (3.2), where the exchange rate is one of two constants depending on the regime and where a first-order autoregressive term captures short run dynamics. Then we extensively discuss the regime process in (3.1). Finally, we briefly address the GARCH-type variance (3.4) and t-distribution (3.5) of the innovation.

As table 3.4 shows, all three exchange rates are characterized by an appreciation and a depreciation regime. Moreover, there is significant first-order autocorrelation for the German mark only.

Despite the minor importance of short-run autocorrelation, all three exchange rates exhibit long-run autocorrelation caused by the long swings, that is, by the persistence of regimes with different means. The high persistence of regimes is represented by the large regime-staying probabilities p_{11} and p_{22}, which all exceed 0.975.

To get a better idea about the degree of persistence that the staying probabilities imply, we first compute the expected duration of a regime r, which is $(1-p_{rr})^{-1}$ (see Hamilton (1989)). The average estimates of p_{11} and p_{22} imply an expected duration of somewhat more than one year for the low mean regime and about two years for the high mean regime.[6]

A second way to examine the persistence of regimes is by inspecting estimated regime probabilities. Following Gray (1996a), we use two types of regime probabilities, namely ex ante and smoothed probabilities. The ex ante probability of a particular regime at time t is the conditional probability that the process was in that regime at time t using only information available to the econometrician at time $t-1$, that is, I_{t-1}. The smoothed regime probability, on the other hand, uses the complete data set I_T, thereby smoothing the ex ante probabilities.[7] Hence,

[6]These durations, which are comparable to the ones in Engel and Hamilton (1990), are remarkably different from the Dewachter (1997) results. For instance, his estimates for the German mark give an expected duration of two and three months instead of years for the low and high mean regime, respectively. Such short durations are difficult to reconcile with his (Wald-based) conclusion that there are long swings. The reason for this result is that he does not take account of short-run autocorrelation. The regimes, which are supposed to model long-run autocorrelation, are then exploited to capture the short-run autocorrelation as well. This leads to unstable regimes and thus to short instead of long swings.

[7]In appendix 3.B we show how to compute the smoothed probabilities in a recursive manner.

Table 3.4: Estimation results

		GERMANY		JAPAN		U.K.	
		RW	RS	RW	RS	RW	RS
Mean of regime	μ_1	0.03 (0.04)	−0.27* (0.09)	0.01 (0.03)	−0.30 (0.15)	0.01 (0.03)	−0.30 * (0.09)
	μ_2		0.15* (0.07)		0.13 (0.07)		0.14 * (0.06)
Autocorr.	θ		0.07* (0.03)		0.04 (0.03)		−0.01 (0.03)
Regime stay prob	p_{11}		0.992 (0.010)		0.976 (0.028)		0.981 (0.021)
	p_{22}		0.996 (0.007)		0.983 (0.019)		0.992 (0.013)
Uncond. variance	σ^2	2.89 (1.08)	3.11 (1.41)	1.82 (0.87)	1.62 (0.84)	2.86 (1.11)	2.81 (1.11)
ARCH	α	0.13* (0.03)	0.14* (0.03)	0.07* (0.02)	0.07* (0.02)	0.11* (0.02)	0.10 * (0.02)
GARCH	β	0.84* (0.04)	0.83* (0.04)	0.92* (0.02)	0.92* (0.02)	0.88* (0.02)	0.89 * (0.02)
T-dist.	ν^{-1}	0.12* (0.03)	0.14* (0.03)	0.20* (0.02)	0.21* (0.02)	0.20* (0.02)	0.22 * (0.03)
Log-likelihood minus RW		-2126 0	-2116 9.34	-2053 0	-2044 8.91	-2068 0	-2062 6.34

Standard errors in parentheses; * is significant at 5% level.
"RW" denotes the random walk, "RS" the regime-switching model of subsection 3.2.1.
We report the inverse of the degrees of freedom of the t-distribution, because testing for conditional normality then boils down to simply testing whether ν^{-1} differs significantly from zero.

the smoothed probability gives the most informative answer to the question which regime the process was likely in at time t.

To illustrate the effect of smoothing the ex ante probabilities, figures 3.1B, 3.2B and 3.3B show the ex ante probabilities of being in the high mean regime for the German mark, Japanese yen and U.K. pound, respectively, while figures 3.1C, 3.2C and 3.3C give the corresponding smoothed probabilities. The ex ante probabilities are, of course, more volatile, in particular the ones for the two European currencies in the first half of the eighties. At that time there were several short periods of depreciation, which were viewed ex ante as indications of regime-

The algorithm is based on Gray (1996b). It links the ex ante probabilities, which are used during estimation (see appendix 3.A), directly to the smoothed probabilities by iterating forward from the ex ante to the smoothed probabilities.

switches. However, they appeared to be only temporary depreciations afterwards, as the dollar continued to strengthen until 1985. Using this information to update the ex ante probabilities smoothes away the temporary deviations and makes the smoothed probabilities much more stable.

The smoothed probabilities in figures 3.1C, 3.2C and 3.3C confirm that regimes are persistent. Moreover, they show that the regime-classification is in general as one would have expected. For instance, the well-known dollar appreciation against the European currencies in the first half of the eighties and the subsequent depreciation against all three currencies are well captured by the regime-switching model.

After this extensive discussion of the regime process, we now briefly address the error variance (3.4) and distribution (3.5). The lower part of table 3.4 contains the estimates. We find that conditional homoskedasticity and conditional normality are strongly rejected. Furthermore, for all three series the results are very robust across the two models, indicating that the variance process is rather independent of the specification of the mean equation.

3.3.4 Diagnostics

To check whether our model sufficiently captures the autocorrelation and conditional heteroskedasticity in the data, we analyze the normalized residuals. Table 3.5 presents tests for autocorrelation and heteroskedasticity in them. From the first-order autocorrelations and the Box-Pierce statistics, we conclude that there is no remaining autocorrelation, at least for the regime-switching model. The random walk, which has no autoregressive term, misses some autocorrelation. Furthermore, the autocorrelation tests for the squared normalized residuals show no reason to extend the variance specifications of the two models.

3.3.5 Forecasting Performance

Knowing that long swings really exist, a natural question is whether this can be exploited to predict future exchange rates better than a random walk. In this subsection we focus on this issue.

We first compare the in-sample and then the out-of-sample forecasts generated by the random walk and the regime-switching model. We examine both

Table 3.5: Diagnostic statistics for normalized residuals and their squares

	GERMANY		JAPAN		U.K.	
	RW	RS	RW	RS	RW	RS
Autocorr. ρ_1	0.10* (0.03)	0.01 (0.03)	0.08* (0.03)	0.01 (0.03)	0.06* (0.03)	0.04 (0.03)
Autocorr. Q_{10}	24.40* [0.01]	6.47 [0.78]	34.11* [0.00]	17.37 [0.07]	16.32 [0.09]	6.37 [0.78]
Autocorr. ρ_1^s	−0.05 (0.03)	−0.05 (0.03)	0.06* (0.03)	0.06* (0.03)	0.03 (0.03)	0.04 (0.03)
Autocorr. Q_{10}^s	16.32 [0.09]	15.87 [0.10]	11.13 [0.35]	11.16 [0.35]	9.31 [0.50]	9.91 [0.45]

Standard errors in parentheses and p-values in square brackets; * is significant at 5% level.
"RW" denotes the random walk, "RS" the regime-switching model of subsection 3.2.1.
The residual is the exchange rate change minus the estimate of its conditional expectation $E_{t-1}\{s_t\}$. The regime probability to integrate out the unobserved regimes in this expectation can be found in appendix 3.A. The residual is normalized using the estimate of its variance $V_{t-1}\{s_t\}$. Note that this variance is not equal to the error variance $V_{t-1}\{\varepsilon_t\}$, since the possibility of regime-switches is an additional source of variance of the residuals besides the error variance.
All autocorrelation statistics have been defined below table 3.2, although the standard error of ρ_1 and the value of Q_{10} are no longer corrected for heteroskedasticity.

point predictions and predictions of the direction of exchange rate changes by comparing the actual (log of the) exchange rate level at some future time τ, S_τ, with the predicted level based on information available at time $t-1$, $\widehat{E}_{t-1}\{S_\tau\}$. For the random walk, this forecast is the previous exchange rate S_{t-1} plus an estimated drift term. For the regime-switching model, $\widehat{E}_{t-1}\{S_\tau\}$ follows from (3.14) in appendix 3.C, after substitution of the estimation results of subsection 3.3.3. The forecasts are computed for three horizons, namely the one-week, which corresponds to the data frequency, the one-quarter (13-week), and the one-year (52-week) horizons.

Starting with the in-sample forecasts, the first often-used forecasting statistics we consider are the root mean squared error (RMSE), which is defined as the square root of $\frac{1}{T}\sum_{t=1}^{T}(S_\tau - \widehat{E}_{t-1}\{S_\tau\})^2$, and the mean absolute error (MAE) $\frac{1}{T}\sum_{t=1}^{T}\left|S_\tau - \widehat{E}_{t-1}\{S_\tau\}\right|$. Table 3.6 shows that the regime-switching model beats the random walk in 12 out of 18 cases, so that there is only a slight preference for our regime-switching model.

Our model, however, clearly outperforms the random walk at predicting the direction of change. In eight out of nine cases the estimated probability of a

Table 3.6: In-sample forecasting statistics

	GERMANY		JAPAN		U.K.	
	RW	RS	RW	RS	RW	RS
Panel A: One-week horizon						
RMSE	1.464	1.458	1.454	1.449	1.459	1.455
MAE	1.095	1.085	1.041	1.033	1.043	1.038
Correct direction	0.527*	0.562*	0.484	0.548*	0.507	0.560 *
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Panel B: One-quarter horizon						
RMSE	5.941	5.959	6.305	6.368	5.974	5.944
MAE	4.814	4.757	4.956	4.916	4.585	4.485
Correct direction	0.530	0.576*	0.539	0.586*	0.492	0.579 *
	(0.045)	(0.041)	(0.047)	(0.038)	(0.046)	(0.039)
Panel C: One-year horizon						
RMSE	12.945	13.487	14.059	14.751	12.891	12.911
MAE	10.585	10.338	11.042	11.581	10.722	10.280
Correct direction	0.534	0.597*	0.609*	0.535	0.480	0.589
	(0.065)	(0.056)	(0.063)	(0.049)	(0.065)	(0.054)

Standard errors in parentheses; * is significantly greater than 0.5 at 5% level.
"RW" denotes the random walk, "RS" the regime-switching model of subsection 3.2.1.
"Correct direction" denotes the fraction of forecasts that yield the correct direction of change of the exchange rate level. For the one-quarter and one-year horizon the standard errors have been corrected for autocorrelation as explained in footnote 8.

correct prediction is higher than for the random walk. In even seven out of nine cases our model predicts the direction of change correctly in significantly more than half of the observations, while for the random walk this happens only once.[8] Apparently, taking account of long swings improves the in-sample forecast quality, particularly regarding the direction of change.

We now turn to the out-of-sample forecasts. We reestimate the two models using only the first three quarters of the sample. Holding the parameters fixed, we then use the 304 observations in the final quarter (from November 1, 1991 to

[8]This conclusion about significance is robust to the autocorrelation originating from the fact that for the one-quarter and one-year horizon the forecast horizon exceeds the one week period between observations. The standard errors of the percentages in table 3.6 are based on the Newey and West (1987) asymptotic covariance matrix. Following West and Cho (1995), we have taken Bartlett weights and have used the same data-dependent automatic lag selection rule. This rule, introduced by Newey and West (1994), has certain asymptotic optimality properties.

Table 3.7: Out-of-sample forecasting statistics

	GERMANY		JAPAN		U.K.	
	RW	RS	RW	RS	RW	RS
Panel A: One-week horizon						
RMSE	1.523	1.526	1.511	1.515	1.465	1.473
MAE	1.133	1.136	1.097	1.099	1.000	1.006
Correct direction	0.512 (0.029)	0.531 (0.029)	0.454 (0.029)	0.484 (0.029)	0.459 (0.029)	0.502 (0.029)
Panel B: One-quarter horizon						
RMSE	5.612	5.680	6.490	6.562	5.638	5.759
MAE	4.589	4.663	5.106	5.026	3.671	3.784
Correct direction	0.438 (0.075)	0.486 (0.076)	0.503 (0.081)	0.545 (0.071)	0.490 (0.074)	0.483 (0.048)
Panel C: One-year horizon						
RMSE	10.151	11.033	12.765	12.803	9.470	9.515
MAE	8.807	9.489	11.059	10.787	6.995	7.297
Correct direction	0.455 (0.101)	0.498 (0.095)	0.605 (0.106)	0.628 (0.080)	0.522 (0.095)	0.522 (0.095)

Standard errors in parentheses; * is significantly greater than 0.5 at the 5% level.
"RW" denotes the random walk, "RS" the regime-switching model of subsection 3.2.1.
The whole series except for the last quarter has been used for estimation, while the last quarter (304 weeks from November 1, 1991 to July 22, 1997) has been used for forecasting. This means that for the one-quarter (one-year) horizon there are 292 (253) comparisons between the actual and predicted values.
"Correct direction" denotes the fraction of forecasts that yield the correct direction of change of the exchange rate level. For the one-quarter and one-year horizon the standard errors have been corrected for autocorrelation as explained in footnote 8.

July 22, 1997) to generate the forecasts $\widehat{E}_{t-1}\{S_\tau\}$.

From table 3.7 we see that the marginal superiority of the regime-switching model in terms of RMSE and MAE has vanished. In only two out of eighteen cases the regime-switching model outperforms the random walk and in the other cases it does worse. This conclusion is also drawn by Engel (1994) using quarterly data on a different out-of-sample period, namely 1986:II to 1991:I. The result is in line with Diebold and Nason (1990), who find in a nonparametric analysis that it is difficult to beat the random walk in point prediction.

Nevertheless, we still see that our model outperforms the random walk at predicting the direction of change, as it does better in seven out of nine cases and does worse only once. This is also concluded by Engel (1994) for his quarterly data set and is supported by our in-sample results.

3.4 Conclusion

The random walk is often used to model exchange rates. We test the validity of that model against the Markov regime-switching model. The latter model generalizes the random walk and explicitly allows for long swings in exchange rates. The central question of the paper is whether such long swings actually exist. We use U.S. dollar exchange rates for the German mark, Japanese yen and U.K. pound from April 1974 to July 1997 to examine this.

In the literature so far (for instance, Engel and Hamilton (1990)), the conclusion is that long swings do exist. However, we demonstrate that the commonly used Wald based tests are not very reliable in the highly nonlinear regime-switching model. Moreover, the random walk cannot be rejected in favor of the long swings when using the more robust likelihood ratio test for similar quarterly data as in Engel and Hamilton (1990).

However, this is not our final conclusion. To be able to test for long swings in higher-frequency data, we introduce a way to extend the basic regime-switching model with GARCH and t-distributed errors to account for conditional heteroskedasticity and extra leptokurtosis. The estimation results provide no evidence for monthly data, but for weekly data we find evidence of long swings in all three exchange rates. Apparently, the sampling frequency matters for tests on long run phenomena such as long swings. Hence, we conclude that long swings are in the data, but that finding them requires fairly high frequency data.

We also analyze whether the existence of long swings can be exploited to forecast exchange rates better than a random walk. As already suggested by Diebold and Nason (1990), beating the random walk in point prediction is difficult. Nevertheless, we find some evidence that the long swings model predicts the direction of change better than the random walk. This gives some out-of-sample support for our conclusion that exchange rates exhibit long swings.

Both empirical results suggest that exchange rates do not behave in a purely random way and that there are occasional changes in regime. Hence, research on the probability of such changes and its variation over time seems promising.

For instance, one can directly analyze the role of policy changes for switches in the exchange rate regime. In this respect, monetary policy announcements may have an effect, as in Kaminsky (1993). Moreover, one can test whether foreign exchange interventions affect the exchange rate regime-switching probabilities, as such interventions may signal changes in future monetary policy (see Loopesko (1984)). One can also include market fundamentals in the regime-switching probabilities, such as the trade balance disequilibrium and deviations from purchasing power parity.

In total, the existence of long swings can stimulate research on the relevance of economic fundamentals for exchange rate determination and, hence, can improve existing theoretical exchange rate models. In Chapter 4 of this book we examine the role of deviations from purchasing power parity as a fundamental for exchange rates. The other suggestions mentioned above are left for future research.

Appendices

3.A Estimation

We estimate the regime-switching model introduced in subsection 3.2.1 by maximum likelihood. In this appendix we derive the likelihood function and show that it has a convenient recursive structure.

To obtain the likelihood function, we first need the density of the exchange rate change at time t conditional on only observable information. Let $p_{t-1}(s_t)$ denote this density evaluated at an exchange rate change equal to s_t.[9] It can be split up as

$$p_{t-1}(s_t) = \sum_{r_t, r_{t-1}=1,2} p_{t-1}(s_t \,|\, r_t, r_{t-1}) \cdot p_{t-1}(r_t, r_{t-1}). \tag{3.6}$$

We now discuss how to compute both terms on the right-hand-side.

The first term, $p_{t-1}(s_t \mid r_t, r_{t-1})$, denotes the density of the exchange rate change at time t evaluated at the value s_t conditional on I_{t-1} and on the current and previous regimes having values r_t and r_{t-1}. This t-density follows from formulas (3.2), (3.4) and (3.5). It is, however, not straightforward how to compute the conditional variance in (3.4), as this requires integrating out the regimes r_{t-1} and r_{t-2} in $\varepsilon_{t-1}^2 = \{s_{t-1} - [\mu_{r_{t-1}} + \theta(s_{t-2} - \mu_{r_{t-2}})]\}^2$. For that, we need $p_{t-1}(r_{t-1}, r_{t-2})$, the conditional probability that the two most recent regimes have values r_{t-1} and r_{t-2}. This probability is crucial, since all regime probabilities in the chapter can be derived from it. Using similar techniques as in Gray (1996a), the following formula shows that this probability has a first-order recursive structure, which simplifies its computation a lot:

$$\begin{aligned} p_{t-1}(r_{t-1}, r_{t-2}) &= p_{t-2}(r_{t-1}, r_{t-2} \,|\, s_{t-1}). \\ &= \frac{p_{t-2}(s_{t-1} \,|\, r_{t-1}, r_{t-2}) \cdot p_{t-2}(r_{t-1}, r_{t-2})}{p_{t-2}(s_{t-1})} \\ &= \frac{p_{t-2}(s_{t-1} \,|\, r_{t-1}, r_{t-2}) \cdot p_{t-2}(r_{t-1} | \, r_{t-2}) \cdot \sum_{r_{t-3}=1,2} p_{t-2}(r_{t-2}, r_{t-3})}{p_{t-2}(s_{t-1})}. \end{aligned} \tag{3.7}$$

[9] We use the same symbol p_{t-1} for several densities (see (3.1) and (3.6)). The specific meaning of p_{t-1} is uniquely determined by the symbols used in its argument. This results in a concise notation, which will prove useful in the remaining part of the chapter.

Hence, the variables to compute $p_{t-1}(r_{t-1}, r_{t-2})$ are its previous values $p_{t-2}(r_{t-2}, r_{t-3})$ for $r_{t-3} = 1, 2$, the constant $p_{t-2}(r_{t-1} | r_{t-2})$ and the previous densities $p_{t-2}(s_{t-1} | r_{t-1}, r_{t-2})$ and $p_{t-2}(s_{t-1})$. This makes the computation of $p_{t-1}(r_{t-1}, r_{t-2})$ a first-order recursive process.

The second term on the right-hand-side of (3.6), $p_{t-1}(r_t, r_{t-1})$, is the conditional probability that the current and previous regimes have values r_t and r_{t-1}, respectively. It can be calculated from

$$p_{t-1}(r_t, r_{t-1}) = p_{t-1}(r_t | r_{t-1}) \cdot \sum_{r_{t-2}=1,2} p_{t-1}(r_{t-1}, r_{t-2}), \tag{3.8}$$

where $p_{t-1}(r_t \mid r_{t-1})$ follows directly from (3.1) and $p_{t-1}(r_{t-1}, r_{t-2})$ is given by (3.7).

Having discussed both terms on the right-hand-side of (3.6), we can now compute the density of interest, $p_{t-1}(s_t)$, being a mixture of four t-densities. This density can then be used to build the sample log-likelihood $\sum_{t=1}^{T} \log(p_{t-1}(s_t))$ with which all parameters in the regime-switching model can be estimated.

From a practical point of view, it is important to realize that the log-likelihood has a second-order recursive structure, similar to that of a standard one-regime AR(1)-GARCH(1,1) model. After all, for (3.8) one needs the constant $p_{t-1}(r_t \mid r_{t-1})$ and the first-order recursive probability $p_{t-1}(r_{t-1}, r_{t-2})$ (see (3.7)) for all eight combinations of (r_t, r_{t-1}, r_{t-2}); density (3.6) can then be computed from (3.8), the previous changes s_{t-1} and s_{t-2}, (3.7) and the previous variance $V_{t-2}\{\varepsilon_{t-1}\}$ in (3.4). This second-order recursiveness of $p_{t-1}(s_t)$ makes the calculation of the sample log-likelihood quite fast. To start up the recursive process, we set the required variables equal to their expectation without conditioning on the information set.

3.B Regime Inference

As stated in footnote 7, the smoothed probability that the regime was r_t at time t, $p_T(r_t)$, can be computed recursively. More generally, any ex post regime probability $p_\tau(r_t)$, for a given future time $\tau \in \{t, t+1, \ldots, T\}$, can be calculated in a recursive manner. This claim, which we prove in this appendix, is based on the following recursive process for the two-regime ex post probability $p_\tau(r_t, r_{t-1})$ starting from the ex ante probability $p_{t-1}(r_t, r_{t-1})$.

We can write $p_\tau(r_t, r_{t-1})$ for the four regime combinations as

$$\begin{aligned} p_\tau(r_t, r_{t-1}) &= p_{\tau-1}(r_t, r_{t-1} \,|\, s_\tau) \\ &= \frac{p_{\tau-1}(s_\tau | r_t, r_{t-1}) \cdot p_{\tau-1}(r_t, r_{t-1})}{\sum_{r_t, r_{t-1}=1,2} p_{\tau-1}(s_\tau | r_t, r_{t-1}) \cdot p_{\tau-1}(r_t, r_{t-1})}. \end{aligned} \tag{3.9}$$

Suppose first that $\tau = t$. Then $p_\tau(r_t, r_{t-1})$ follows directly from (3.9), as $p_{\tau-1}(r_t, r_{t-1})$ and $p_{\tau-1}(s_\tau | r_t, r_{t-1})$ are known from the estimation process (see appendix 3.A).

Hence, let us suppose from now on that $\tau > t$. The computation of (3.9) requires two inputs. The first one is the previous ex post probability $p_{\tau-1}(r_t, r_{t-1})$, which is known from the previous recursion for all combinations of r_t and r_{t-1}. The second ingredient of (3.9) is the density $p_{\tau-1}(s_\tau | r_t, r_{t-1})$ for all regime outcomes. Its computation requires a number of steps. We first write it as

$$p_{\tau-1}(s_\tau | r_t, r_{t-1}) = \sum_{r_\tau, r_{\tau-1}=1,2} p_{\tau-1}(s_\tau | r_\tau, r_{\tau-1}) \cdot p_{\tau-1}(r_\tau, r_{\tau-1} | r_t, r_{t-1}), \tag{3.10}$$

where we use that the conditional distribution of s_τ given $r_\tau, r_{\tau-1}$ does not depend on the earlier regimes r_t and r_{t-1}. This formula itself has two ingredients. The first one is the density $p_{\tau-1}(s_\tau | r_\tau, r_{\tau-1})$ for all regime combinations, which is known from the estimation process.

The second term needed in (3.10) is the $(\tau - t)$-period-ahead regime-switching probability $p_{\tau-1}(r_\tau, r_{\tau-1} | r_t, r_{t-1})$ for all regime combinations. Once it has been computed, it should be saved, since it will be needed in the next recursive step. Making use of the Markov structure of the regime process, it can be written in terms of $(\tau - 1 - t)$-period-ahead switching probabilities:

$$p_{\tau-1}(r_\tau, r_{\tau-1} | r_t, r_{t-1}) = \sum_{r_{\tau-1}, r_{\tau-2}=1,2} p_{\tau-1}(r_\tau, r_{\tau-1} | r_{\tau-1}, r_{\tau-2}) \cdot p_{\tau-1}(r_{\tau-1}, r_{\tau-2} | r_t, r_{t-1}). \tag{3.11}$$

Again, we have two ingredients. First, we need $p_{\tau-1}(r_\tau, r_{\tau-1} | r_{\tau-1}, r_{\tau-2})$ for all regime combinations. Due to the Markov property of the regime process, this switching probability does not depend on $r_{\tau-2}$. It equals

$$p_{\tau-1}(r_\tau, r_{\tau-1} | r_{\tau-1}, r_{\tau-2}) = p_{\tau-1}(r_\tau | r_{\tau-1}), \tag{3.12}$$

which is constant and follows from (3.1).

The second ingredient of (3.11) is $p_{\tau-1}(r_{\tau-1}, r_{\tau-2}|r_t, r_{t-1})$ for all regime combinations:

$$\begin{aligned} p_{\tau-1}(r_{\tau-1}, r_{\tau-2}|r_t, r_{t-1}) &= p_{\tau-2}(r_{\tau-1}, r_{\tau-2}|r_t, r_{t-1}, s_{\tau-1}) \\ &= \frac{p_{\tau-2}(s_{\tau-1}|r_{\tau-1}, r_{\tau-2}) \cdot p_{\tau-2}(r_{\tau-1}, r_{\tau-2}|r_t, r_{t-1})}{\sum_{r_{\tau-1}, r_{\tau-2}=1,2} p_{\tau-2}(s_{\tau-1}|r_{\tau-1}, r_{\tau-2}) \cdot p_{\tau-2}(r_{\tau-1}, r_{\tau-2}|r_t, r_{t-1})}, \end{aligned} \tag{3.13}$$

where we use that the conditional density of $s_{\tau-1}$ is independent of the previous regimes r_t and r_{t-1} once $r_{\tau-1}$ and $r_{\tau-2}$ are given. We have two ingredients. First, the conditional density $p_{\tau-2}(s_{\tau-1}|r_{\tau-1}, r_{\tau-2})$ for all regime combinations. It is known from the estimation process. Second, we need the $(\tau-1-t)$-period-ahead switching probability $p_{\tau-2}(r_{\tau-1}, r_{\tau-2}|r_t, r_{t-1})$ for all regime combinations. This one was saved during the previous recursion, if $\tau > t+1$. If $\tau = t+1$, it equals one.

This completes the algorithm to compute (3.10), which is the second ingredient of (3.9). For each recursion one has to compute (3.13), use it together with (3.12) to compute (3.11) and use this to compute (3.10). Using this in (3.9) yields the ex post probability $p_\tau(r_t, r_{t-1})$ from $p_{\tau-1}(r_t, r_{t-1})$. Therefore, starting from the ex ante probability $p_{t-1}(r_t, r_{t-1})$ one can recursively compute the ex post probability $p_\tau(r_t, r_{t-1})$ and eventually the probability of interest $p_\tau(r_t)$.

3.C Forecasting

Subsection 3.3.5 deals with forecasting exchange rate levels S_τ at time $t-1$, where $\tau \geq t$. This appendix explains how to compute these forecasts.

As usual, we first decompose the exchange rate forecast as

$$E_{t-1}\{S_\tau\} = S_{t-1} + \sum_{i=t}^{\tau} E_{t-1}\{s_i\}. \tag{3.14}$$

To calculate $E_{t-1}\{s_i\}$, we rewrite s_i by repeated substitution of lags of (3.2) for the lagged changes. Since the innovations have zero expectation, this yields

$$E_{t-1}\{s_i\} = \sum_{r_i, r_{t-1}=1,2} p_{t-1}(r_i, r_{t-1}) \cdot \left(\mu_{r_i} + \theta^{i-(t-1)}(s_{t-1} - \mu_{r_{t-1}})\right), \tag{3.15}$$

where

$$p_{t-1}(r_i, r_{t-1}) = p_{t-1}(r_{t-1}) \cdot p_{t-1}(r_i \,|r_{t-1}), \tag{3.16}$$

where $p_{t-1}(r_{t-1})$ follows after summation of $p_{t-1}(r_{t-1}, r_{t-2})$ in (3.7) over r_{t-2}.

To compute the multi-period-ahead switching probability $p_{t-1}(r_i \,|\, r_{t-1})$ on the right-hand-side of (3.16), we first form the one-period-ahead Markov transition matrix M:

$$M = \begin{bmatrix} p(r_t{=}1 \,|\, r_{t-1}{=}1) & 1 - p(r_t{=}2 \,|\, r_{t-1}{=}2) \\ 1 - p(r_t{=}1 \,|\, r_{t-1}{=}1) & p(r_t{=}2 \,|\, r_{t-1}{=}2) \end{bmatrix}, \tag{3.17}$$

where its elements follow from (3.1). The theory of Markov processes for multi-period-ahead switching probabilities then implies that

$$p_{t-1}(r_i \,|\, r_{t-1}) = \left(M^{i-(t-1)}\right)_{r_i r_{t-1}}. \tag{3.18}$$

Having explained how to calculate (3.16), we can now compute (3.15). Computation of (3.15) for all i and substitution in (3.14) then gives the forecast of interest $E_{t-1}\{S_\tau\}$.

Chapter 4

Purchasing Power Parity: Evidence from a New Test

Most economists intuitively consider purchasing power parity (PPP) to be true. Nevertheless, the empirical literature is not very supportive of PPP. In this chapter, however, we find evidence in favor of PPP using a new test approach. It is based on a Markov regime-switching model for the exchange rate, because earlier papers have shown that this model seems more realistic than the popular random walk. We allow for PPP by making the regime-switching probabilities depend on the PPP deviation. Our second result is that PPP disequilibria have become shorter-lived for some exchange rates, which may be due to an increase in trade openness of the countries involved.

4.1 Introduction

Purchasing Power Parity (PPP) is one of the oldest theories in international economics. Moreover, it is a building block of other theories, for instance, theories of exchange rate determination. PPP is commonly used as a long-run concept in relative terms, stating that in the long-run the (nominal) exchange rate is proportional to the ratio of the two countries' price levels, that is, the PPP exchange rate. Long-run relative PPP is also the version of PPP we use.

This chapter examines the empirical validity of PPP. This is relevant in several respects. From a theoretical point of view, the validity of PPP is important for research on theories of exchange rate determination, particularly concerning the long-run. Practically, PPP can provide a target exchange rate for monetary authorities, which they can use for foreign exchange interventions, among other things. In addition, the long-term behavior of exchange rates relative to prices is relevant for international firms. They have to decide upon foreign direct investments and therefore require reliable forecasts of the real value of the long-lasting income stream generated by the investment projects. Taking account of long-run PPP, if valid, may help improve the long-run exchange rate forecasts they need.

Most economists intuitively consider the PPP hypothesis to be true. Moreover, time plots of exchange rates PPP rates support it. For instance, figures 4.1A, 4.2A and 4.3A in section 4.3 plot the U.S. dollar price of one German mark, Japanese yen and U.K. pound, respectively, and the corresponding PPP rates from April 1974 to July 1997; the figures suggest a long-run comovement of exchange rates and PPP rates (details on the construction of the PPP rates will be given in subsection 4.3.1).

Quite surprisingly, however, the voluminous existing empirical literature is not very supportive of PPP, that is, the null of no PPP is not often rejected. The main contribution of this chapter is that we reject the absence of PPP for all three of the world's most important exchange rates mentioned above.

The reason behind this remarkable difference is that we use a new test approach. It is based on a Markov regime-switching model (see Hamilton (1989)) that uses two regimes for the mean exchange rate change to allow for long swings in exchange rates. This model seems more realistic than the popular random walk (with drift), as we argue below. We show that PPP holds in the long swings model if a swing is likely to end when the PPP disequilibrium becomes large and if the next swing governs the exchange rate back to its PPP equilibrium. Hence, to test for PPP we examine whether these conditions are valid.

Given the evidence in favor of PPP, it is natural to examine what the economic mechanism behind PPP is. The common argument for PPP is that goods arbitrage equalizes prices in the same currency across countries. Because it is commonly believed that goods markets have become more integrated, making arbitrage easier, it is interesting to examine whether PPP disequilibria have become shorter-lived, which is the second and final purpose of the chapter. We conclude that they have for the German mark and the U.K. pound, but not for the Japanese yen. This may indeed be explained by changes in trade openness, as we find that both European economies have become much more open, while Japan is still relatively closed.

The remaining part of this introduction presents a brief overview of the existing empirical literature on PPP and explains the contribution of this chapter in more detail.

In the literature so far, many authors have examined PPP (see Froot and Rogoff (1996) for a detailed overview). They usually concentrate on the real exchange rate and employ unit-root tests to examine the null that the real exchange rate follows a random walk against the alternative of stationarity, that is, PPP. Early studies, such as Meese and Rogoff (1988) and Mark (1990), use post-Bretton-Woods time series data and find no evidence of PPP. This may, of course, be caused by the absence of PPP. For example, some goods are not traded across countries, so that the goods arbitrage argument for price equalization and hence PPP (see above) no longer holds. However, the insignificant results may also be due to a lack of power of the tests or because the random walk setting is inappropriate for exchange rates. After all, the null of no PPP is in fact a joint null hypothesis of the absence of PPP and the validity of the random walk model, so that the outcome of the test can be affected by the random walk assumption.

As Frankel (1986, 1990) shows, a potential reason for the lack of power is that the post-Bretton-Woods period may be too short to contain enough episodes of divergence from and reversion to PPP, because PPP disequilibria may dampen very slowly. This suggestion has resulted in two approaches to increase the power of the test. First, Frankel (1986) and Abuaf and Jorion (1990), among others, use very long time series, often extending to a century. They indeed find evidence in favor of PPP. There is, however, some concern with these results, since the long-horizon time series blend fixed and floating rate data, and it is well-known that real exchange rates behave very differently under different exchange rate regimes (see Mussa (1986)). This is why we use only post-Bretton-Woods data.

A second way to gain power, while using only floating data, is to analyze a panel of many countries. Two notable studies in this field are Frankel and Rose (1996) and Papell (1997), which both find evidence in favor of PPP. Recently, however, O'Connell (1998) reports that the panel evidence disappears if one takes account of the strong cross-sectional dependence in real exchange rates. This argument does not apply to our results, as we analyze three exchange rates univariately.

In summary, Frankel's (1986, 1990) suggestion that PPP tests may lack power because of the use of relatively short post-Bretton-Woods time series has not resulted in conclusive evidence of PPP, despite the enormous number of studies motivated by this suggestion.

In the present chapter we start from a different point of view. As mentioned above, the lack of evidence for PPP in the literature may be due to a lack of power of the unit-root tests or because the random walk setting is inappropriate for exchange rates. We reduce both potential problems simultaneously by using a more general model and another test.

The model we propose is the Markov regime-switching model. It generalizes the random walk, as the latter is a special case in which the regimes coincide.

The regime-switching model seems more realistic than the random walk, both from an economic and an empirical point of view. After all, the regime-switching model allows for some changes in the exchange rate generating process, which, according to the Lucas (1976) critique, may result from changes in economic policy. For instance, regarding monetary policy, Kaminsky (1993) shows theoretically that a change from a contractionary to an expansionary monetary policy increases the exchange rate depreciation and that this makes a regime-switching model more appropriate. Regarding international policy coordination, the 1985 Plaza agreement (the G-5 countries announced to try to bring about a U.S. dollar depreciation after the sharp dollar appreciation during the five years before) seemed to have an effect on the exchange rate generating process, as the dollar depreciated strongly from 1985 to 1987. Both examples show that policy shifts can lead to changes in the trend of exchange rates and thus to long swings. Such swings are a systematic part of the regime-switching model, not of the random walk. Moreover, there is empirical evidence that such swings indeed exist. For instance, see Engel and Hamilton (1990), Engel (1994) and Chapter 3, which reject the random walk in favor of the regime-switching model. Hence, the regime-switching model seems a more appropriate setting than the random walk to test for PPP.

Testing for PPP within a regime-switching framework for the (nominal) exchange rate is not standard, as existing regime-switching models do not take account of PPP. They often assume that the probability of switching to the, say, depreciation regime is constant over time. However, PPP implies that such a switch becomes more likely when the currency is overvalued regarding its PPP value. Thus, to develop a test for PPP, we first extend the basic regime-switching model by allowing the regime-switching probabilities to depend on the PPP deviation.[1] We then derive three parameter restrictions under which the extended model implies PPP, and we test the joint validity of these restrictions. Because this test clearly supports PPP, the reason for the insignificant results from the unit-root random walk tests in the existing literature is not the absence of PPP, but rather a lack of power or a misspecification of the random walk model.

In the next section, we define the regime-switching model. In section 4.3 we describe the data and present our empirical results. Section 4.4 concludes.

4.2 Regime-Switching Model

In this section we develop the regime-switching model that we use to answer the two questions of the chapter, namely whether PPP holds and, if so, whether PPP disequilibria have become shorter-lived. We first set out the basic principles in an intuitive way. In subsections 4.2.1, 4.2.2 and 4.2.3 we then formally develop the model in three stages, where each stage generalizes the previous one.

The basic idea of our approach is that exchange rates exhibit two types of long swings, for instance, an appreciation and a depreciation swing. A random process governs the switches between the swings (or regimes). This regime-switching process is crucial, as the variants of the model in subsections 4.2.1, 4.2.2 and 4.2.3 differ with respect to this process only.

In the simplest model, see subsection 4.2.1, the probability of switching from one regime to the other is constant over time. Hence, the level of the exchange rate is irrelevant for the switching probabilities.

In subsection 4.2.2 we generalize this assumption, because it contradicts with PPP. After all, PPP implies that switches to the, say, depreciation regime are more likely when the currency is overvalued regarding PPP; the exchange rate is pulled towards its PPP equilibrium. To allow for this pull, we let the regime-switching probabilities depend on the PPP deviation. Interestingly, it appears that long-run relative PPP holds, if the pull is present and if the swing-specific appreciation and depreciation are strong enough compared to the PPP rate change

[1] Time-varying switching probabilities are also useful when modeling switches between recessions and recoveries in business cycles; see Durland and McCurdy (1994), Filardo (1994) and Ghysels (1994).

(so that the exchange rate is able to return to its PPP rate after an under or overvaluation, respectively). Hence, to test for PPP, we can test whether these three conditions are fulfilled.

To answer the second question of the chapter, about the duration of PPP disequilibria, we need one further generalization. It is based on the idea that PPP ensures that the long swings are around the PPP equilibrium. Therefore, PPP disequilibria become shorter-lived if the long swings around PPP get shorter. In subsection 4.2.3 we thus allow for a change in the duration of the swings and describe how to test whether this change is negative.

In the remaining part of this section, we formally work out the intuition just given.

4.2.1 Regime-Switching Model Without PPP

The regime-switching model without PPP is based on Hamilton (1989). The main difference with the basic Hamilton model is that we allow for conditional heteroskedasticity, which is present in the weekly data we use in the empirical application.

To describe the model, we need the following notation. As in Chapters 2 and 3, let S_t denote the logarithm of the nominal spot exchange rate at time t, that is, the domestic currency price of one unit of foreign currency. We again concentrate on the exchange rate change $s_t = 100(S_t - S_{t-1})$, so that s_t is the percentage depreciation of the domestic currency from time $t-1$ to t.

The regime-switching model consists of four elements. Two of them, the regime process and the mean equation, are crucial for interpreting our empirical results, as they are directly related to the exchange rate swings. The other two, the variance and distribution, will be discussed at the end of this subsection.

The regime process is based on two (unobservable) regimes. As in Chapter 3, let $r_t \in \{1, 2\}$ denote the regime at time t. Within this regime, the mean exchange rate change is μ_{r_t}, which we assume to be constant over time. Across regimes, however, the means are allowed to differ, and we identify the first regime as the low mean regime: $\mu_1 \leq \mu_2$. This provides the basis for the swings. After all, being in the first and then in the second regime for a while leads to a period of appreciation followed by depreciation, that is, to swings in the exchange rate.

Whether swings are long or not depends on the regime staying probabilities.

Let $p_{t-1}(r_t|\tilde{r}_{t-1}) = p(r_t|I_{t-1}, \tilde{r}_{t-1})$ denote the probability of going to regime r_t at time t conditional on the information set of the data generating process, which consists of two parts. The first part, I_{t-1}, denotes the information that is observed by the econometrician; in this subsection I_{t-1} consists of $(s_{t-1}, s_{t-2}, \dots)$. The second part, $\tilde{r}_{t-1}$, is the regime path $(r_{t-1}, r_{t-2}, \dots)$, which is not observed by the econometrician. Note that we use the subscript $t-1$ below an operator (probability, expectation or variance) as short-hand notation for conditioning on I_{t-1}.

As in the Hamilton (1989) model, we assume in this subsection that r_t follows a first-order Markov process with constant staying probabilities, so that

$$p_{t-1}(r_t\,|\,\tilde{r}_{t-1}) = p(r_t\,|\,r_{t-1}) = \begin{cases} p_{11} & \text{if } r_t = r_{t-1} = 1 \\ p_{22} & \text{if } r_t = r_{t-1} = 2. \end{cases} \tag{4.1}$$

Hence, if p_{11} and p_{22} are high, regimes are persistent and exchange rate swings are long. Note that in (4.1) the conditional probability that the current regime is low or high depends on the past (I_{t-1} and $\tilde{r}_{t-1}$) only through the most recent regime r_{t-1}. This assumption represents the only difference between the current model and its generalizations in the next two subsections.

Whereas persistence in mean regimes is supposed to take account of the long swings, or "long-run autocorrelation", there may still be short-run dynamics within a mean regime. In the conditional mean specification we take account of this "short-run autocorrelation" by using one autoregressive term, as it is generally believed that the short-run autocorrelation in exchange rates is small (see West and Cho (1995)):

$$s_t = \mu_{r_t} + \theta(s_{t-1} - \mu_{r_{t-1}}) + \varepsilon_t, \tag{4.2}$$

where the conditional expectation of the innovation is $E_{t-1}\{\varepsilon_t|\tilde{r}_t\} = 0$.

The regime process and conditional mean just specified are the most important elements of the model. For a complete model specification, however, we also have to define the two other elements, namely the conditional variance and distribution. This is done in the remaining part of this subsection.

When specifying the conditional variance of the error term in (4.2), $V_{t-1}\{\varepsilon_t|\tilde{r}_t\}$, we take account of the conditional heteroskedasticity in the weekly data that we use in the empirical application. We use the following generalized autoregressive conditional heteroskedasticity (GARCH) type model (see Bollerslev, Chou and Kroner (1992) for an overview of GARCH in standard, one-regime models):

$$V_{t-1}\{\varepsilon_t|\widetilde{r}_t\} = V_{t-1}\{\varepsilon_t\} = \omega + \alpha E_{t-1}\{\varepsilon_{t-1}^2\} + \beta V_{t-2}\{\varepsilon_{t-1}\}, \tag{4.3}$$

with the usual GARCH restrictions $\omega > 0$ and α, $\beta \geq 0$ to ensure $V_{t-1}\{\varepsilon_t\} > 0$ for all t. We also assume $\alpha + \beta < 1$, so that the unconditional variance is $\sigma^2 = \frac{\omega}{1-\alpha-\beta}$. Note that we set $V_{t-1}\{\varepsilon_t|\widetilde{r}_t\}$ equal to its expectation conditional on only observable information I_{t-1}, that is, $V_{t-1}\{\varepsilon_t\}$. This is only for the sake of estimation simplicity.[2] We admit that it is a restriction. However, the purpose of the variance specification is only to make the PPP results, which we are mainly interested in, robust to conditional heteroskedasticity. Subsection 4.3.4 shows that (4.3) is sufficient for that.

The fourth and final element of our model, the conditional error distribution, is specified by a t-distribution, which is often used to allow for extra leptokurtosis (see Bollerslev, Chou and Kroner (1992)). It has ν degrees of freedom, zero mean, and variance $V_{t-1}\{\varepsilon_t\}$:

$$\varepsilon_t \,|\, I_{t-1}, \widetilde{r}_t \quad \sim \quad t\,(\nu\ ,\ 0\ ,\ V_{t-1}\{\varepsilon_t\})\,. \tag{4.4}$$

This completes the regime-switching model without PPP; it is given by (4.1) to (4.4).

4.2.2 Regime-Switching Model With PPP

In this subsection we extend the model of the previous subsection, so as to be able to test whether (long-run relative) PPP holds. We first examine the implications of PPP for the model and show why the model needs some extension. The required extension turns out to deal with the regime-staying probabilities in

[2] If we had not set $V_{t-1}\{\varepsilon_t|\widetilde{r}_t\} = V_{t-1}\{\varepsilon_t\}$, the variance would have been $V_{t-1}\{\varepsilon_t|\widetilde{r}_t\} = \omega + \alpha\varepsilon_{t-1}^2 + \beta V_{t-2}\{\varepsilon_{t-1}|\widetilde{r}_{t-1}\}$ and would have depended on the entire regime path up to time $t-1$. After all, r_{t-1} and r_{t-2} would have affected the variance through the surprise term ε_{t-1}^2, which is $\{s_{t-1} - [\mu_{r_{t-1}} + \theta(s_{t-2} - \mu_{r_{t-2}})]\}^2$ expressed in the conditioning variables, and earlier regimes would have affected the variance through the earlier surprise terms implicitly present in the lagged variance term. This would have rendered estimation intractable, since the number of possible regime paths is enormous and all regime paths have to be integrated out for estimation, as they are not observed. Specification (4.3) circumvents this problem by directly averaging out the regimes r_{t-1} and r_{t-2} in the source of the path-dependence, ε_{t-1}^2. The basic idea of this technique originates from Gray (1996a) and is further discussed in Chapter 3.

(4.1) only. Having described the implications of PPP, we then show that these implications also imply PPP, so that a test on their joint validity delivers a test for PPP. Finally, we give the test statistic that we will use in the empirical study.

According to PPP, the deviation of the exchange rate S_t from the PPP exchange rate S_t^{ppp}, being the (logarithm of the) home price level over the foreign price level, is constant in the long-run. Therefore, if the current PPP deviation is higher (lower) than this constant, the PPP deviation is expected to fall (rise) in the long run.

In the regime-switching model, this has three implications. First, to make a fall in the PPP deviation possible, the expected change μ_1 in the low mean regime must, of course, be smaller than the expected depreciation of the PPP rate, μ_{ppp}, say. Similarly, to make a rise in the PPP deviation possible, μ_2 must exceed μ_{ppp}; this is the second implication of PPP.

The third implication concerns the regime process. In the model without PPP, the regime-staying probabilities (4.1) are constant over time. This is unrealistic if PPP holds. After all, if the current PPP deviation is, say, higher than the long-run constant, the probability of going to the low mean regime increases, so as to swing the process back into the direction of its PPP equilibrium. Hence, a large PPP deviation increases the probability of staying in the low mean regime but decreases the probability of staying in the high mean regime.

To model this dependence of the regime-staying probabilities on the PPP deviation, $S_{t-1} - S_{t-1}^{ppp}$, we use a logit specification for simplicity:

$$p_{t-1}(r_t \,|\, \tilde{r}_{t-1}) = \begin{cases} \Lambda(\delta_1 + \delta_{ppp}(S_{t-1} - S_{t-1}^{ppp})) & \text{if } r_t = r_{t-1} = 1 \\ \Lambda(\delta_2 - \delta_{ppp}(S_{t-1} - S_{t-1}^{ppp})) & \text{if } r_t = r_{t-1} = 2, \end{cases} \tag{4.5}$$

where $\Lambda(.)$ is the standard logistic distribution function.[3] For parsimony, we restrict the effect of the PPP deviation to be the same (in absolute sense) for both probabilities, so that a single parameter, δ_{ppp}, captures the effect of PPP. This parameter is positive if PPP holds, and it measures the strength with which the exchange rate is pulled towards PPP equilibrium. Note that for $\delta_{ppp} = 0$ the staying probabilities are simply $\Lambda(\delta_1)$ and $\Lambda(\delta_2)$, which correspond to p_{11} and p_{22} in (4.1), respectively.

[3] As opposed to the model without PPP, the information set of the econometrician, I_{t-1}, now consists of the previous exchange rate and PPP rate levels. As before, the information of the data generating process also contains the regime path.

So far, we have concentrated on the implications of PPP for the regime-switching model: $\delta_{ppp} > 0$ is the necessary pull towards equilibrium, and $\mu_1 < \mu_{ppp} < \mu_2$ is necessary for PPP because otherwise the exchange rate will move away from PPP even if $\delta_{ppp} > 0$. To get a test for PPP, however, we need to know what these three restrictions tell us about PPP. In appendix 4.A we show through simulations that the restrictions imply that PPP holds.[4] Hence, one can test the null of no PPP by testing the joint null of $\mu_1 \geq \mu_{ppp}$ or $\mu_{ppp} \geq \mu_2$ or $\delta_{ppp} \leq 0$, which is the complement of the three restrictions mentioned above. Given the existing literature, as described in the introduction, this is a new way to test for PPP.

In the remaining part of this subsection, we develop the test statistic we use in subsection 4.3.2 of our empirical study. We assume for simplicity that the expected PPP depreciation, μ_{ppp}, is given. This makes the null only depend on the vector $\pi = (\mu_1, \mu_2, \delta_{ppp})'$ of parameters of the regime-switching model. Since the null consists of several inequality constraints on π, we define our test statistic along the lines of Kodde and Palm (1986). That is, we use the distance from the data, represented by the maximum likelihood (ML) estimate $\widehat{\pi}$ of π, to the closest feasible point under the null (see appendix 4.C for a description of the ML estimation procedure). More formally, our PPP test statistic is

$$\widehat{\Pi} = \min_{\pi \in H_0} (\widehat{\pi} - \pi)' \widehat{V}\{\widehat{\pi}\}^{-1} (\widehat{\pi} - \pi), \tag{4.6}$$

where H_0 is the set of feasible vectors π under the null, and $\widehat{V}\{\widehat{\pi}\}$ is the ML estimate for the variance of $\widehat{\pi}$.

Definition (4.6) shows that $\widehat{\Pi} \geq 0$ and that only points $\widehat{\pi} \notin H_0$ lead to $\widehat{\Pi} > 0$. To determine whether a realization of $\widehat{\Pi}$ is sufficiently positive to reject the null, we need the distribution of $\widehat{\Pi}$ under the null. However, we cannot use the theory in Kodde and Palm (1986) for that. After all, under the null of no PPP, the PPP deviation $S_{t-1} - S_{t-1}^{ppp}$ in the regime-staying probabilities (4.5) is non-stationary, making the distribution of $\widehat{\pi}$ and hence $\widehat{\Pi}$ potentially non-standard. Therefore, we simulate the null distribution of $\widehat{\Pi}$. Appendix 4.B describes the simulation procedure that we use for our empirical study.

[4]More formally, we show that the mean and variance of the PPP deviation are constant in the long-run and that the respective constants are independent of the current situation. This is what one usually means with the verbal statement that according to PPP the PPP deviation is constant in the long-run, because the latter interpretation is unreasonably strict.

4.2.3 Duration of PPP Disequilibria

Having extended the basic regime-switching model with the allowance for PPP, we need one further extension to be able to examine the second issue of the chapter, namely whether PPP disequilibria, being the difference between PPP deviations and their long run constant value, have become shorter-lived. Of course, this question is only relevant if PPP holds. Therefore, the current subsection is conditional on this. As in the previous subsection, we first extend the model to allow for a change in the duration of PPP disequilibria, and then we present the test that we use in the empirical study.

In the regime-switching model with PPP, the duration of PPP disequilibria changes if the duration of the swings around PPP changes. Since the latter depends on the intercepts in the regime-staying probabilities (4.5), we allow for a break in these intercepts:

$$p_{t-1}(r_t\,|\,\tilde{r}_{t-1}) = \begin{cases} \Lambda(\delta_{10} + \delta_{ppp}(S_{t-1} - S_{t-1}^{ppp}) + \delta_{11} d_{t-1}) & \text{if } r_t = r_{t-1} = 1 \\ \Lambda(\delta_{20} - \delta_{ppp}(S_{t-1} - S_{t-1}^{ppp}) + \delta_{21} d_{t-1}) & \text{if } r_t = r_{t-1} = 2, \end{cases} \tag{4.7}$$

where d_t is one if time t is after the break date and zero otherwise.

To complete (4.7), we have to choose the break date. Of course, such a choice is rather ad hoc. However, from an economic point of view, the Louvre accord of February 22, 1987 is an interesting break date. After all, the Louvre accord exactly aimed at stabilizing exchange rates by introducing target zones, so as to prevent the long dollar swings of the years before. Therefore, negative values for δ_{11} and δ_{21} in (4.7) represent that PPP disequilibria have become shorter-lived after the Louvre accord.

To test whether δ_{11} and δ_{21} are negative, we use their ML-based t-values. These t-values have standard (normal) limit distributions, because $S_{t-1} - S_{t-1}^{ppp}$ is stationary in case of PPP. Hence, one can use standard inference. Subsection 4.3.3 presents the results.

4.3 Empirical Results

In this section we use the regime-switching model of section 4.2 to answer the two questions of this chapter, namely whether relative PPP holds in the long-run and whether PPP disequilibria have become shorter-lived. First, we describe the data. Then, in subsection 4.3.2 we test for PPP and in 4.3.3 we examine the duration of PPP disequilibria. In subsection 4.3.4 we check the specification of our model. In the last subsection, we analyze whether taking account of PPP leads to better exchange rate forecasts than the simple random walk model.

4.3.1 Data

We use three U.S. dollar exchange rates, namely, the dollar vis-à-vis the German mark, the Japanese yen and the British pound. These exchange rates have been chosen because of their important role on foreign exchange markets and because they behave relatively independently, for instance, compared to dollar-EMS exchange rates. We take weekly instead of monthly or quarterly data, because Chapter 3 reports for the same series and model strategy that only weekly data yield enough observations to significantly distinguish a long swings process from a random walk, and because our central parameter δ_{ppp}, measuring the strength with which swings are pulled towards PPP, is only identified if there are swings. The data set contains 1,216 observations for the percentage dollar depreciations s_t from April 2, 1974 to July 22, 1997.

To construct the PPP exchange rates S_t^{ppp}, we follow most of the literature by using consumer price indices from the IMF International Financial Statistics.[5] They have been obtained from Datastream, just as the exchange rates. In the remaining part of this subsection, we analyze the characteristics of the three exchange rates and PPP rates and use them to motivate our model specification empirically.

In panel A of figures 4.1, 4.2 and 4.3, we show the behavior of the three actual and PPP exchange rates over the sample period (in U.S. dollars, not in logarithms). At first sight, exchange rates seem to be characterized by long swings. This impression is formally tested for the same data in Chapter 3, where we find that long swings are indeed part of the exchange rate generating process. This

[5] We use a linear interpolation to generate weekly PPP rates from the available monthly rates. The interpolation method one chooses is practically irrelevant for the results, because PPP rates are much more stable than actual exchange rates.
For illustrative convenience, we add a constant to the price index ratios such that the average PPP deviation is zero. This only affects the estimates for the constant terms in the logit specifications (4.5) and (4.7). Hence, it does not affect any of our tests of interest.

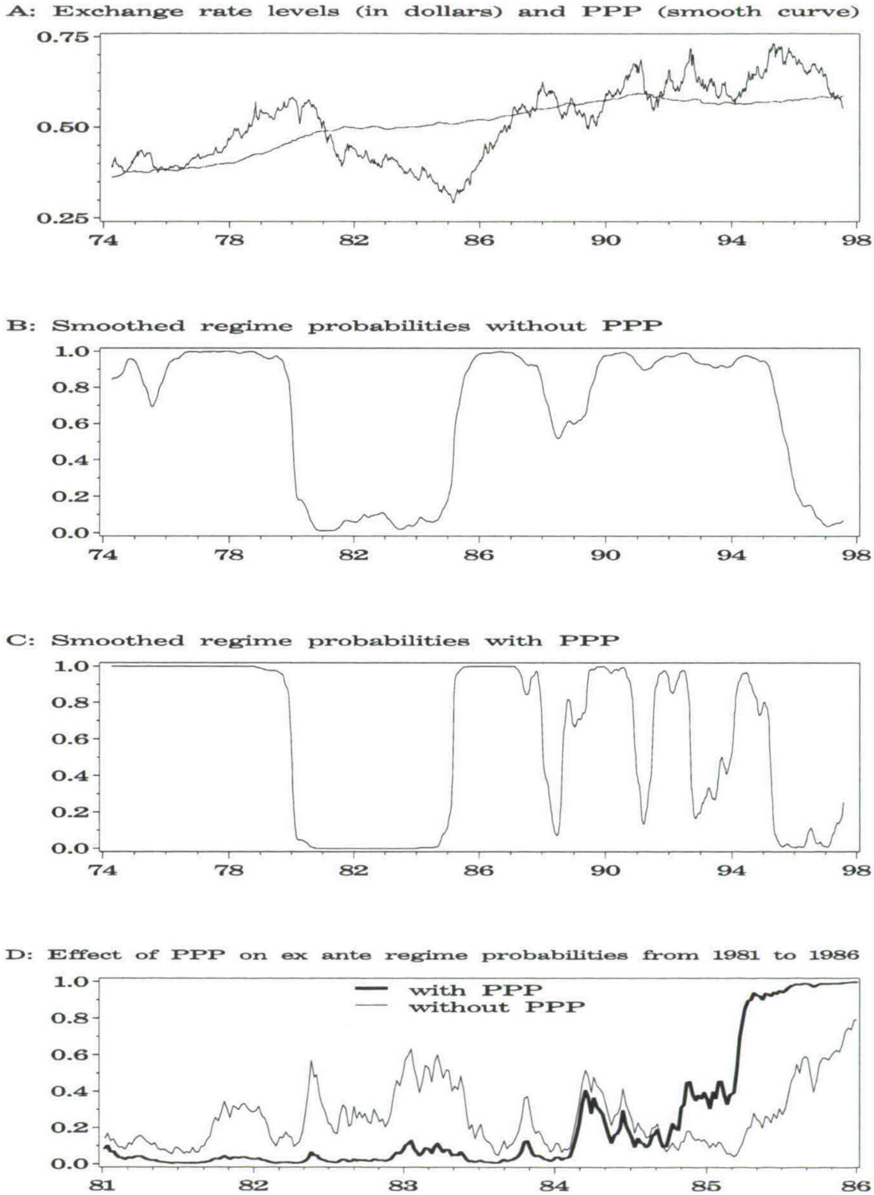


Figure 4.1: German mark over the sample period April 1974 to July 1997

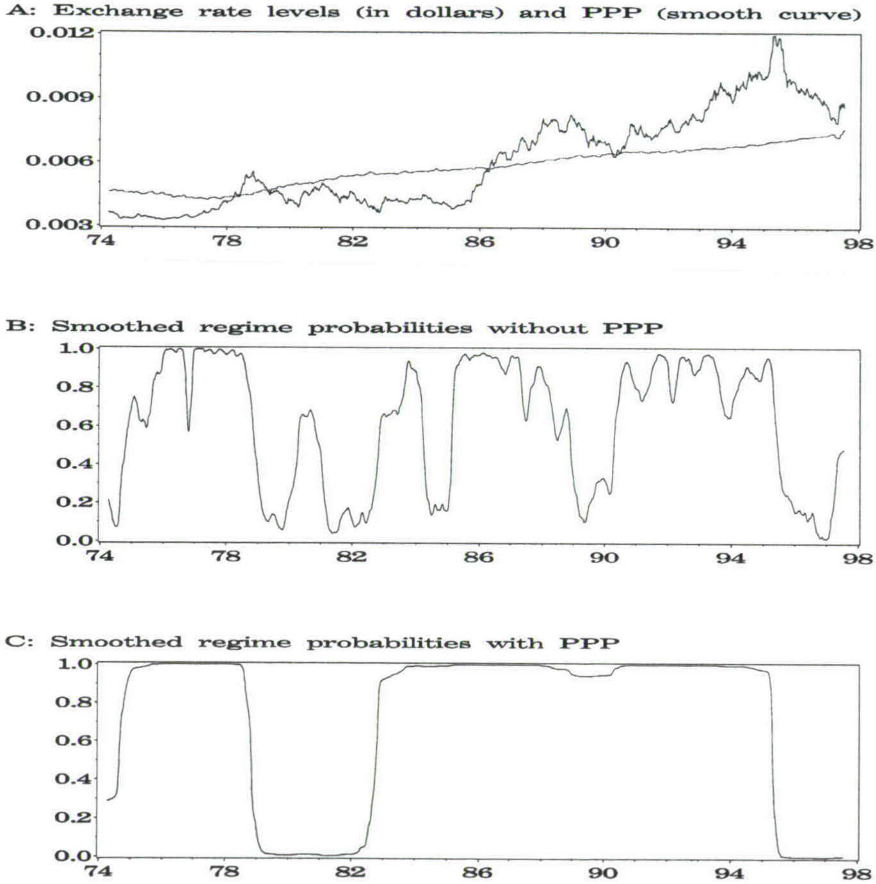


Figure 4.2: Japanese yen over the sample period April 1974 to July 1997

motivates the use of a regime-switching model empirically (see the introduction for theoretical motivations).

The figures also suggest that exchange rates swing around PPP and that the swings are likely to end when the deviation from the PPP rate is large. Therefore, it seems useful to let the regime-switching probabilities depend on the PPP deviation, as our model does.

Finally, we see from the plots that the swings for the two European currencies

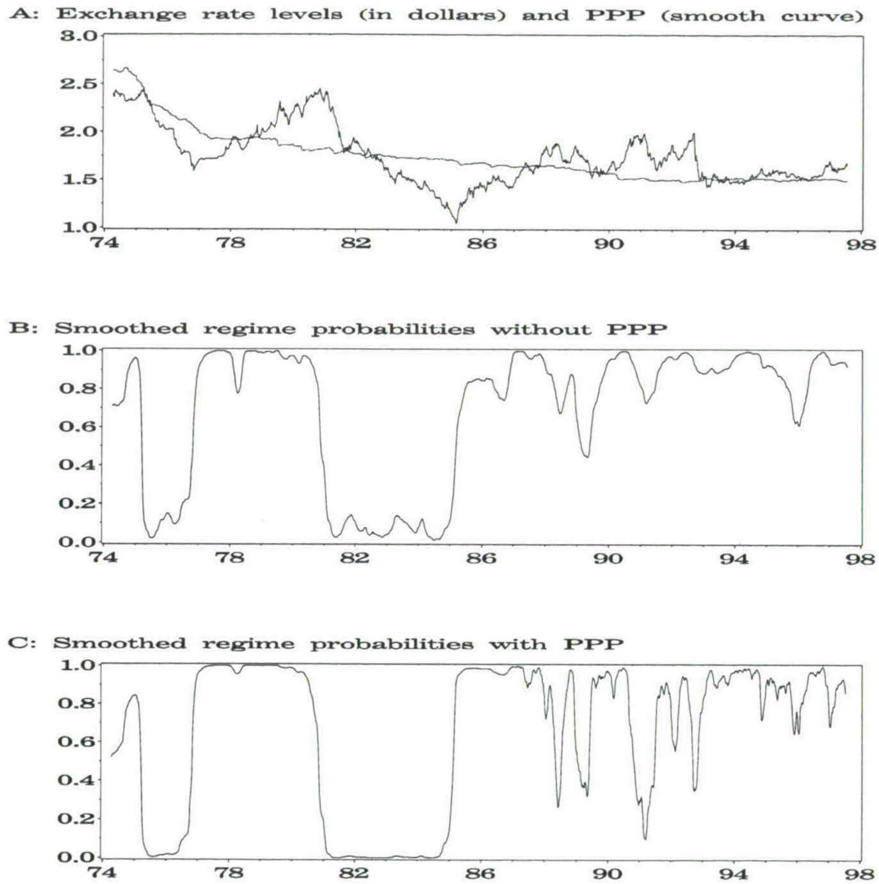


Figure 4.3: U.K. pound over the sample period April 1974 to July 1997

seem to be shorter in the second half of the sample. This shows that there may well have been a break in the duration of the swings. Our model allows for that.

In table 4.1 we report some descriptive statistics of the three exchange rate changes. There is significant first-order autocorrelation in the weekly German mark changes (we always use a significance level of 5%). Estimates for higher-order autocorrelations are not reported separately, but are combined in Box-Pierce type statistics $\widetilde{Q}_{10}$. They show that higher-order autoregressive terms in

Table 4.1: Moments of exchange rate changes and autocorrelation tests

	GERMANY	JAPAN	U.K.
Mean	0.03	0.07	−0.03
Variance	2.14	2.11	2.13
Skewness	−0.14	0.53	−0.40
Excess Kurtosis	1.70	2.01	3.00
Autocorr. ρ_1	0.07* (0.03)	0.05 (0.04)	0.04 (0.04)
Autocorr. $\widetilde{Q}_{10}$	14.07 [0.17]	22.57* [0.01]	6.05 [0.81]
Autocorr. squares ρ_1^s	0.11* (0.03)	0.20* (0.03)	0.20 * (0.03)
Autocorr. squares Q_{10}^s	57.60* [0.00]	92.03* [0.00]	151.82 * [0.00]

Standard errors in parentheses and p-values in square brackets; * is significant at 5% level.
The first-order autocorrelation, ρ_1, has been estimated as the slope coefficient in a regression of the change, s_t, on the first lagged change, s_{t-1}, and a constant. The standard errors are based on White's (1980) heteroskedasticity-consistent asymptotic covariance matrix.
$\widetilde{Q}_{10}$ denotes a modified Box-Pierce type statistic that combines the first ten autocorrelations. Following Pagan and Schwert (1990), it is defined as the sum of the first ten squared normalized autocorrelation estimates, where the normalizing factors are the heteroskedasticity-consistent standard errors of the autocorrelation estimates. $\widetilde{Q}_{10}$ is asymptotically χ^2_{10} distributed.
The first-order autocorrelation in the squared changes, ρ_1^s, and the Box-Pierce type statistic Q_{10}^s are similarly defined, although without the correction for heteroskedasticity.

the mean equation (4.2) are unnecessary.

Table 4.1 also presents two autocorrelation tests for the squared exchange rate changes. Both tests point at conditional heteroskedasticity for all three series. This is why we have extended the basic Hamilton (1989) regime-switching model with GARCH specification (4.3) for the conditional error variance.

4.3.2 Does Relative Purchasing Power Parity Hold in the Long Run?

In this subsection we use the theory of 4.2.2 to answer the central question of the chapter. That is, we compute the PPP statistic $\widehat{\Pi}$ in (4.6) to test the null of no PPP using the estimation results for the regime-switching model with PPP but without the post-Louvre dummy.[6]

The results for $\widehat{\Pi}$ [and its p-value] are: 5.46 [0.01] for Germany, 3.14 [0.05] for Japan, and 4.61 [0.01] for the U.K.[7] Hence, we find evidence in favor of long-run relative PPP for all three U.S. dollar exchange rates over the post-Bretton-Woods period of floating. Given that the existing literature is not very supportive of PPP, it is remarkable that we find such conclusive evidence with our new test. This shows that the results in the literature so far are not due to the absence of PPP. Apparently, the unit-root type tests in the random walk setting that are commonly used are not the most appropriate ways to test for PPP; the tests are not powerful enough given the relatively short post-Bretton-Woods data period, or the random walk is not the most appropriate model for exchange rates.

The existence of PPP has important implications for the exchange rate swings. More specifically, exchange rate swings are more likely to end when the PPP deviation is large (see subsection 4.2.2). To illustrate this effect, figure 4.1D contains the ex ante probability of being in the high mean regime for Germany for the regime-switching models with and without PPP from 1981 to the beginning of 1986.[8] According to the model without PPP, the temporary upward moves between 1982 and 1985 are interpreted as signs of regime shifts. However, some weeks later, it appears that there has been no such shift, and the ex ante probabilities become low again. The regime probabilities of the model with PPP are much less affected by the temporary upward moves in the early eighties. However, when the PPP deviation gets larger, their effect increases.

[6]Because the estimation results for the model without the post-Louvre dummy are similar to the ones for the model with the post-Louvre dummy (to be discussed below), we do not report the estimation results of the former model to save space.

[7]Appendix 4.B describes how we have simulated the p-values. It also argues that these p-values are conservative, that is, they are likely somewhat higher than the true p-values. However, in our case this is no problem, as the reported p-values are already low.
At first sight, it may be surprising that the p-values are so low given that the $\widehat{\Pi}$ are not very large. After all, for a test on a single one-sided restriction in a standard setting of all stationary variables, the asymptotic 5% critical value is $1.65^2 = 2.71$, where 1.65 is the 5% quantile of the standard normal distribution, and this value generally increases when non-stationary variables are involved. However, in our case the alternative hypothesis of PPP consists of three instead of a single one-sided restriction, and this has a negative effect on the critical value.

[8]The ex ante regime probability for time t is defined as the conditional probability that the process is in a particular regime at time t using only the information set of the econometrician at time $t-1$, that is, I_{t-1} (see Gray (1996a)).

4.3.3 Have PPP Disequilibria Become Shorter-Lived?

From the previous subsection, we know that the PPP deviation is constant in the long run. In the short run, however, there are considerable periods in which the PPP deviation is different from this constant. In the current subsection we examine whether such PPP disequilibria have become shorter-lived, the second theme of the chapter. We use the theory of subsection 4.2.3.

As argued in 4.2.3, we test for a change in the duration of PPP disequilibria by testing whether the swings around the PPP rate get shorter after the Louvre accord in 1987. More formally, we test whether the parameters δ_{11} and δ_{21} for the post-Louvre dummy in (4.7) are negative. The results follow from table 4.2, which presents the estimates of all parameters in our model, as well as two benchmark models, namely the random walk and the regime-switching model without PPP. The table demonstrates that PPP disequilibria have become shorter-lived for the two European currencies, as three out of four coefficients for the post-Louvre dummy are significantly negative. However, we find no evidence of shorter PPP disequilibria for the yen.

The shorter duration of PPP disequilibria for the European currencies may be caused by attempts to stabilize exchange rates, such as the Louvre accord. Another reason may be the increased openness of countries. For instance, the ratio of total trade over output, which is often used as a measure for openness, has increased over our period of observation 1974-1997 from 0.42 to 0.58 for Germany and from 0.41 to 0.64 for the U.K.[9]

The shorter duration of PPP disequilibria is graphically illustrated by figures 4.1C and 4.3C, which plot the smoothed regime probabilities of being in the high mean regime for Germany and the U.K., respectively.[10] For both exchange rates we observe more, but much shorter swings after 1987, so that the exchange rates

[9]The underlying total trade (exports plus imports) and output figures are from the OECD Main Economic Indicators in Datastream.

[10]The difference between the smoothed regime probability at time t and the ex ante probability, as defined in footnote 8, is that the former probability uses the complete data set I_T instead of only I_{t-1}, thereby smoothing the ex ante probabilities. Hence, the smoothed regime probability gives the most informative answer to the question which regime the process was likely in at time t. In appendix 4.D we show how to compute the smoothed probabilities in a recursive manner. The algorithm is based on Gray (1996b). It links the ex ante probabilities, which are used during estimation, directly to the smoothed probabilities by iterating forward from the ex ante to the smoothed probabilities.

Table 4.2: Estimation results

		GERMANY			JAPAN			U.K.		
		RW	noPPP	PPP	RW	noPPP	PPP	RW	noPPP	PPP
Mean of regime	μ_1	0.03 (0.04)	-0.27* (0.09)	-0.29* (0.07)	0.01 (0.03)	-0.30 (0.15)	-0.36* (0.09)	0.01 (0.03)	-0.30* (0.09)	-0.31* (0.10)
	μ_2		0.15* (0.07)	0.20* (0.05)		0.13 (0.07)	0.09* (0.03)		0.14* (0.06)	0.16* (0.05)
Autocorr.	θ		0.07* (0.03)	0.06* (0.03)		0.04 (0.03)	0.05 (0.03)		-0.01 (0.03)	-0.01 (0.03)
Regime stay prob	p_{11}		0.992 (0.010)			0.976 (0.028)			0.981 (0.021)	
	p_{22}		0.996 (0.007)			0.983 (0.019)			0.992 (0.013)	
Logit intercept	δ_{10}			10.98* (3.12)			8.08* (2.56)			5.85* (1.34)
post-Louvre	δ_{11}			-7.77* (3.09)			0			-4.44* (1.65)
	δ_{20}			10.28* (2.82)			9.10* (2.96)			5.63* (1.19)
	δ_{21}			-4.93* (2.18)			0			-1.31 (1.34)
PPP deviation	δ_{ppp}			17.18 (7.29)			13.53 (7.63)			8.13 (4.23)
Uncond. variance	σ^2	2.89 (1.08)	3.11 (1.41)	3.07 (1.35)	1.82 (0.87)	1.62 (0.84)	1.80 (0.90)	2.86 (1.11)	2.81 (1.11)	2.73 (1.06)
ARCH	α	0.13* (0.03)	0.14* (0.03)	0.15* (0.03)	0.07* (0.02)	0.07* (0.02)	0.07* (0.02)	0.11* (0.02)	0.10* (0.02)	0.10* (0.02)
GARCH	β	0.84* (0.04)	0.83* (0.04)	0.82* (0.04)	0.92* (0.02)	0.92* (0.02)	0.92* (0.02)	0.88* (0.02)	0.89* (0.02)	0.89* (0.02)
T-dist.	ν^{-1}	0.12* (0.03)	0.14* (0.03)	0.14* (0.03)	0.20* (0.02)	0.21* (0.02)	0.21* (0.02)	0.20* (0.02)	0.22* (0.03)	0.21* (0.03)
Log-likelihood minus RW		-2126 0	-2116 9.34	-2110 15.82	-2053 0	-2044 8.91	-2043 10.41	-2068 0	-2062 6.34	-2057 11.28

Standard errors in parentheses; * is significant at 5% level.
"RW" denotes the random walk, "noPPP" ("PPP") the regime-switching model without (with) allowance for PPP (see (4.1) and (4.7), respectively.)
Because of our evidence in favor of PPP, the PPP deviation $S_{t-1}-S^{ppp}_{t-1}$ in (4.7) is stationary. Therefore, the t-values for all parameters except δ_{ppp} have the standard (normal) asymptotic distribution, so that one can use standard inference. For δ_{ppp} the t-value may well have a non-standard limit distribution, so that we do not know for sure whether the estimates in the table are significant.
The two zero entries in table 4.2 for Japan indicate that we have to impose $\delta_{21}=\delta_{22}=0$ to achieve convergence. This restriction is realistic, as figure 4.2A shows no signs of a structural break in the yen-dollar swings after the Louvre accord.
We present the inverse of the degrees of freedom of the t-distribution, because testing for conditional normality then boils down to simply testing whether ν^{-1} differs significantly from zero.

do not move far away from their PPP rates. For the U.K. the increased stability makes it even difficult to classify the observations after 1987 into one particular regime, which leads to the fairly unstable smoothed regime plot.

The second conclusion mentioned above, the lack of evidence of shorter PPP disequilibria for Japan, is in contrast with the conclusions for the two European currencies. This is, however, not surprising, because the Japanese economy is still quite closed, at least compared to Germany and the U.K., as the trade/output ratio has increased from 0.17 in 1974 to only 0.25 in 1997. This makes Japanese economic policy more independent, so that PPP disequilibria can be more persistent.

4.3.4 Diagnostics

The results of the two previous subsections are all based on a regime-switching model. In this subsection, we check the specification of that model in two ways, namely by testing whether the model takes account of all autocorrelation and conditional heteroskedasticity in the data. We use the normalized residuals for that.

Table 4.3 presents the test results, not only for our preferred model, but also for the two benchmark models introduced before. From the first-order autocorrelations and the Box-Pierce statistics Q_{10}, we conclude that there is no remaining autocorrelation, at least for the two regime-switching models. Furthermore, the first-order autocorrelation and the aggregate autocorrelation test Q^s_{10} for the squared normalized residuals show no reason to extend the variance specifications of the three models.

4.3.5 Forecasting Performance

Knowing that PPP holds and that PPP disequilibria have become shorter-lived for Germany and the U.K., we now examine whether this can be exploited to predict future exchange rates better than a random walk.

We first compare the in-sample and then the out-of-sample forecasts generated by the random walk and the regime-switching model with and without PPP. We examine both point predictions and predictions of the direction of the exchange rate change by comparing the actual (logarithm of the) exchange rate level at

Table 4.3: Diagnostic statistics for normalized residuals and their squares

	GERMANY			JAPAN			U.K.		
	RW	noPPP	PPP	RW	noPPP	PPP	RW	noPPP	PPP
Autocorr. ρ_1	0.10*	0.01	0.01	0.08*	0.01	0.01	0.06*	0.04	0.03
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Autocorr. Q_{10}	24.40*	6.47	6.53	34.11*	17.37	18.86*	16.32	6.37	5.87
	[0.01]	[0.78]	[0.77]	[0.00]	[0.07]	[0.04]	[0.09]	[0.78]	[0.83]
Autocorr. ρ_1^s	-0.05	-0.05	-0.06	0.06*	0.06*	0.06*	0.03	0.04	0.04
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Autocorr. Q_{10}^s	16.32	15.87	17.75	11.13	11.16	10.99	9.31	9.91	10.21
	[0.09]	[0.10]	[0.06]	[0.35]	[0.35]	[0.36]	[0.50]	[0.45]	[0.42]

Standard errors in parentheses and p-values in square brackets; * is significant at 5% level.
"RW" denotes the random walk, "noPPP" ("PPP") the regime-switching model without (with) allowance for PPP (see (4.1) and (4.7), respectively.)
The residual is the exchange rate change minus the estimate of its conditional expectation $E_{t-1}\{s_t\}$. The regime probability to integrate out the unobserved regimes in this expectation can be found in appendix 4.C. The residual is normalized by its variance, $V_{t-1}\{s_t\}$. Note that it is not equal to the error variance $V_{t-1}\{\varepsilon_t\}$, since the possibility of regime-switches is an additional source of variance of the residuals besides the one represented by the error term.
All autocorrelation statistics have been defined below table 4.1, although the standard error of ρ_1 and the value of Q_{10} are no longer corrected for heteroskedasticity.

some future time τ, S_τ, with the predicted level based on information available at time t–1, $\widehat{E}_{t-1}\{S_\tau\}$. For the random walk, this forecast is the previous exchange rate S_{t-1} plus an estimated drift term. For the regime-switching model, $\widehat{E}_{t-1}\{S_\tau\}$ follows from (4.17) in appendix 4.E, after substitution of the estimation results of table 4.2. The forecasts are computed for three horizons, namely the one-week horizon, which corresponds to the data frequency, the one-quarter (13-week), and the one-year (52-week) horizons.

Starting with the in-sample forecasts, the first often-used forecasting statistics we consider are the root mean squared error (RMSE), which is the square root of $\frac{1}{T}\sum_{t=1}^{T}(S_\tau-\widehat{E}_{t-1}\{S_\tau\})^2$, and the mean absolute error (MAE) $\frac{1}{T}\sum_{t=1}^{T}\left|S_\tau-\widehat{E}_{t-1}\{S_\tau\}\right|$. Table 4.4 presents their values. They show that our regime-switching model with PPP-based switching probabilities beats both the random walk and the regime-switching model without PPP in 14 out of 18 cases. The four cases where it is not the best model are all for the yen. This currency has only very few swings within our sample, so that it is not surprising that regime-switching forecasts and forecasts from a random walk are of about equal quality.

Although there is a slight preference for our regime-switching model according

Table 4.4: In-sample forecasting statistics

	GERMANY			JAPAN			U.K.		
	RW	noPPP	PPP	RW	noPPP	PPP	RW	noPPP	PPP
Panel A: One-week horizon									
RMSE	1.464	1.458	1.448	1.454	1.449	1.452	1.459	1.455	1.445
MAE	1.095	1.085	1.080	1.041	1.033	1.033	1.043	1.038	1.034
Correct direction	0.527*	0.562*	0.562*	0.484	0.548*	0.552*	0.507	0.560*	0.561*
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Panel B: One-quarter horizon									
RMSE	5.941	5.959	5.522	6.305	6.368	6.369	5.974	5.944	5.482
MAE	4.814	4.757	4.347	4.956	4.916	4.914	4.585	4.485	4.217
Correct direction	0.530	0.576*	0.687*	0.539	0.586*	0.591*	0.492	0.579*	0.647*
	(0.045)	(0.041)	(0.036)	(0.047)	(0.038)	(0.045)	(0.046)	(0.039)	(0.038)
Panel C: One-year horizon									
RMSE	12.945	13.487	11.035	14.059	14.751	14.280	12.891	12.911	9.724
MAE	10.585	10.338	8.411	11.042	11.581	11.210	10.722	10.280	7.668
Correct direction	0.534	0.597*	0.736*	0.609*	0.535	0.648*	0.480	0.589	0.767*
	(0.065)	(0.056)	(0.049)	(0.063)	(0.049)	(0.057)	(0.065)	(0.054)	(0.046)

Standard errors in parentheses; * is significantly greater than 0.5 at 5% level.
"RW" denotes the random walk, "noPPP" ("PPP") the regime-switching model without (with) allowance for PPP (see (4.1) and (4.7), respectively).
"Correct direction" denotes the fraction of forecasts that yield the correct direction of change of the exchange rate level. For the one-quarter and one-year horizon the standard errors have been corrected for autocorrelation as explained in footnote 11.

to the RMSE and MAE, our model clearly outperforms the other models at predicting the direction of change. In all nine cases the estimated probability of a correct prediction is higher than for the two other models. Moreover, in all cases our model predicts the direction of change correctly in significantly more than half of the observations, while for the random walk this happens in only one case.[11] This outperformance can be attributed to two features. First, the long swings improve the forecast quality, as the regime-switching model without PPP already outperforms the random walk in eight cases. Second, the allowance

[11]These conclusions about significance are robust to the autocorrelation originating from the fact that for the one-quarter and one-year horizon the forecast horizon exceeds the one week period between observations. The standard errors of the percentages are based on the Newey and West (1987) asymptotic covariance matrix. Following West and Cho (1995), we have taken Bartlett weights and have used the same data-dependent automatic lag selection rule. This rule, introduced by Newey and West (1994), has certain asymptotic optimality properties.

Table 4.5: Out-of-sample forecasting statistics

	GERMANY			JAPAN			U.K.		
	RW	noPPP	PPP	RW	noPPP	PPP	RW	noPPP	PPP
Panel A: One-week horizon									
RMSE	1.523	1.526	1.526	1.511	1.515	1.544	1.465	1.473	1.463
MAE	1.133	1.136	1.139	1.097	1.099	1.098	1.000	1.006	1.012
Correct direction	0.512	0.531	0.502	0.454	0.484	0.539	0.459	0.502	0.502
	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)
Panel B: One-quarter horizon									
RMSE	5.612	5.680	5.810	6.490	6.562	7.643	5.638	5.759	5.506
MAE	4.589	4.663	4.575	5.106	5.026	5.935	3.671	3.784	3.928
Correct direction	0.438	0.486	0.599	0.503	0.545	0.575	0.490	0.483	0.594
	(0.075)	(0.076)	(0.072)	(0.081)	(0.071)	(0.075)	(0.074)	(0.048)	(0.071)
Panel C: One-year horizon									
RMSE	10.151	11.033	13.301	12.765	12.803	20.394	9.470	9.515	8.253
MAE	8.807	9.489	11.528	11.059	10.787	17.234	6.995	7.297	7.339
Correct direction	0.455	0.498	0.573	0.605	0.628	0.553	0.522	0.522	0.684*
	(0.101)	(0.095)	(0.095)	(0.106)	(0.080)	(0.101)	(0.095)	(0.095)	(0.084)

Standard errors in parentheses and p-values in square brackets; * is significantly greater than 0.5 at 5% level.
"RW" denotes the random walk, "noPPP" ("PPP") the regime-switching model without (with) allowance for PPP (see (4.1) and (4.7), respectively).
The whole series except for the last quarter has been used for estimation, while the last quarter (304 weeks from November 1, 1991 to July 22, 1997) has been used for forecasting. This means that for the one-quarter (year) horizon there are 292 (253) comparisons between the actual and predicted values.
"Correct direction" denotes the fraction of forecasts that yield the correct direction of change of the exchange rate level. For the one-quarter and one-year horizon the standard errors have been corrected for autocorrelation as explained in footnote 11.

for PPP in the switching probabilities leads to additional predictive power. This holds particularly at long horizons, which is in line with the fact that PPP is a long-run phenomenon.

We now turn to the out-of-sample forecasts. We reestimate the two models using only the first three quarters of the sample. Holding the parameters fixed, we then use the 304 observations in the final quarter (from November 1, 1991 to July 22, 1997) to generate the forecasts $\widehat{E}_{t-1}\{S_\tau\}$.

From table 4.5 we see that the superiority of our regime-switching model with PPP-based switching probabilities has vanished, at least in terms of RMSE and MAE. In only four out of eighteen cases our model outperforms both other models (in the other cases it does at least worse than the random walk). Especially for Japan our model seems to do badly. This has the same reason as given above: the swings in the yen-dollar rate are so long that there are only three switches in the in-sample period (see figure 4.2C). Because such switches are crucial for identifying the switching-probability parameters, the parameter estimates differ substantially from the ones based on the complete sample. Hence, more data are needed for the yen to give our model a fair chance.

Concentrating on the European currencies only, the fact that our model does not outperform the random walk may, again, be due to the rather low number of regime-switches in the in-sample period. However, it may also indicate that it is indeed difficult to beat the random walk in point prediction, as Diebold and Nason (1990) conclude from their nonparametric analysis.

Notwithstanding this result, we do find that our model outperforms the random walk at predicting the direction of change, particularly at longer horizons. The outperformance is partly due to the long swings, as the regime-switching model without PPP does already better than the random walk. Engel (1994) also reports this latter result and finds that the outperformance is particularly at the short-run. Our model with PPP-based switching probabilities, however, does particularly well at longer horizons, likely because PPP is a long-run phenomenon. The in-sample forecasting results led to the same conclusion.

4.4 Conclusion

In this chapter we analyze the popular hypothesis of purchasing power parity (PPP), more specifically, long-run relative PPP. The main contribution of the chapter is that we find evidence in favor of PPP for the world's three main U.S. dollar exchange rates over the post-Bretton-Woods period, namely the dollar vis-à-vis the German mark, Japanese yen and U.K. pound. This likely implies that PPP also holds for several other currencies closely linked to them, such as the French franc and the Dutch guilder, which closely follow the German mark.

Our evidence of PPP is remarkable, because the extensive existing literature is not very supportive of PPP. The reason for this difference is that we use a new test approach. It is based on a regime-switching model for the nominal exchange rate, in which the regime-switching probabilities depend on the PPP deviation.

We show that under three simple restrictions, this model yields PPP.

Although the validity of PPP is interesting in itself, it also stimulates other research. For instance, PPP is a building block of many traditional structural exchange rate models, so that its validity underscores their usefulness for long-run exchange rate determination. Moreover, our result helps the development of new exchange rate models. As an example, consider the theory of exchange rate bubbles, where the exchange rate can diverge from its equilibrium value as determined by an economic model (see De Grauwe (1990)). The dependence of the long swings on the PPP deviation in our model indicates that bubbles tend to burst when the deviation from PPP becomes large. Hence, including PPP in bubble theories seems fruitful.

Given the existence of PPP, we can also examine the reasons behind PPP. Our results support the view that goods arbitrage is one of the factors underlying PPP, as we find that PPP disequilibria have become shorter-lived for those countries (Germany and the U.K.) that have the largest increase in trade over the period of observation.

Our third result is that the existence of long-run PPP makes the predictions of the direction of exchange rate changes generated by our model better than those from the popular random walk model, particularly many periods ahead. The relative performance in point prediction, however, is not yet clear, because the post-Bretton-Woods data period is too short compared to the length of the swings to get sufficiently accurate in-sample estimates for the regime-switching parameters. This problem can be reduced by pooling several exchange rate series in a panel data set and then imposing some cross-sectional parameter restrictions to increase estimation accuracy. This is left for future research.

Our model can be extended in various ways. Firstly, other variables such as forward rates can be included in the mean equation to improve exchange rate forecasts. Secondly, variables as the trade balance or monetary policy indicators may be informative about regime-switches, so that it may prove useful to include them besides the PPP deviation in the regime-switching probabilities.

Although we have shown that regime-switching models can provide a framework for testing long-run PPP, they may also be useful to test other long-run relationships. This is due to the interesting feature that a process that is non-stationary within regimes (in our case the nominal exchange rate process) can be transformed into a stationary process (real exchange rate) by letting the regime-switching probabilities depend on a second variable (the PPP deviation). This idea embodies a contribution of our work to the theoretical regime-switching literature. It can be used, for instance, to test the long-run quantity theory of money, stating that the price level is proportional to the money supply in the long term. Hence, regime-switching models may offer an alternative for unit-root tests that are commonly employed to test for long-run relations. These issues are left for future research.

Appendices

4.A Three Parameter Restrictions Imply that PPP Holds

In 4.2.2 we have claimed that $\mu_1 < \mu_{ppp} < \mu_2$ and $\delta_{ppp} > 0$ imply PPP. This appendix verifies that. For that, we first express PPP in more formal terms.

In words, the theory of (long-run relative) PPP states that the PPP deviation is constant in the long-run. Of course, constancy is a very strict requirement. One usually means that the mean and variance of the PPP deviation are constant in the long-run and that the respective constants are independent of the current situation. We follow this interpretation. Therefore, PPP formally means that both $E_{t-1}\{S_\tau - S_\tau^{ppp}|\tilde{r}_{t-1}\}$ and $V_{t-1}\{S_\tau - S_\tau^{ppp}|\tilde{r}_{t-1}\}$ converge (for $\tau \to \infty$) to a limit that is independent of the conditioning information I_{t-1} and $\tilde{r}_{t-1}$, that is, the paths of exchange rates, PPP rates and regimes up to time $t-1$.

Because we have not yet succeeded to derive a formal proof for our claim that $\mu_1 < \mu_{ppp} < \mu_2$ and $\delta_{ppp} > 0$ imply PPP, we use a simulation experiment to show that it is very likely true.[12] This experiment consists of two parts. First, we demonstrate for one particular value of the initial exchange rate level S_{t-1} and the initial PPP deviation $S_{t-1} - S_{t-1}^{ppp}$, which are the only relevant parts of I_{t-1} in our simulation experiment, that under the three constraints $E_{t-1}\{S_\tau - S_\tau^{ppp}|\tilde{r}_{t-1}\}$ and $V_{t-1}\{S_\tau - S_\tau^{ppp}|\tilde{r}_{t-1}\}$ converge to a limit that is independent of the initial regime r_{t-1}, the only relevant part of $\tilde{r}_{t-1}$. In the second part, we show that the two limits are also independent of the initial exchange rate and PPP deviation.

To verify the first part of our claim, we simulate both moments $E_{t-1}\{S_\tau - S_\tau^{ppp}|\tilde{r}_{t-1}\}$ and $V_{t-1}\{S_\tau - S_\tau^{ppp}|\tilde{r}_{t-1}\}$ for horizons one to 2,000 time periods and check our claim graphically. For that, we generate two data sets of 100,000 series of 2,000 future PPP deviations $S_\tau - S_\tau^{ppp}$. All series of both data sets start from

[12] The reported simulation results are based on the following parameter values for the regime-switching exchange rate process: $\mu_1 = -0.2$, $\mu_2 = 0.2$, $\rho = 0$, $\omega = 2.5$, $\alpha = 0$, $\beta = 0$, $\nu = \infty$ and $(\delta_{10}, \delta_{11}, \delta_{20}, \delta_{21}, \delta_{ppp}) = (7, 0, 7, 0, 10)$ (the symmetry is only for the ease of interpretation). Although our model in section 4.2 leaves the PPP exchange rate process unspecified, we have to assume some process for the simulation study. For simplicity, we assume a random walk process: $s_\tau^{ppp} = 100(S_\tau^{ppp} - S_{\tau-1}^{ppp}) = \mu_{ppp} + \eta_\tau$, where $\mu_{ppp} = 0.05$ and η_τ is standard normally distributed. We have tried various other combinations, each satisfying $\mu_1 < \mu_{ppp} < \mu_2$ and $\delta_{ppp} > 0$, and all yield essentially the same results.

$S_{t-1}=0$ and $S_{t-1}-S_{t-1}^{ppp}=0$, and all series within the first (second) data set are based on r_{t-1} equal to one (two). The simulated value of $E_{t-1}\{S_\tau - S_\tau^{ppp}|\tilde{r}_{t-1}\}$ ($V_{t-1}\{S_\tau - S_\tau^{ppp}|\tilde{r}_{t-1}\}$) is defined as the mean (variance) of the 100,000 drawings of the future PPP deviation. In figures 4.4A and B, the two curves labeled $\delta_{ppp} > 0$ plot these simulated mean and variance, respectively, for all horizons. It is clear that $E_{t-1}\{S_\tau - S_\tau^{ppp}|\tilde{r}_{t-1}\}$ and $V_{t-1}\{S_\tau - S_\tau^{ppp}|\tilde{r}_{t-1}\}$ indeed converge to a limit that does not depend on r_{t-1} and, therefore, not on $\tilde{r}_{t-1}$.

For comparison, figures 4.4A and B also contain the simulated moments in case the parameters do not satisfy the joint restrictions $\mu_1 < \mu_{ppp} < \mu_2$ and $\delta_{ppp} > 0$. Since it is obvious that $E_{t-1}\{S_\tau - S_\tau^{ppp}|\tilde{r}_{t-1}\}$ and $V_{t-1}\{S_\tau - S_\tau^{ppp}|\tilde{r}_{t-1}\}$ do not converge if the interval of regime-specific exchange rate means does not contain the PPP trend, $\mu_{ppp} \notin (\mu_1, \mu_2)$, we only concentrate on $\delta_{ppp} \leq 0$.[13] Suppose first that $\delta_{ppp} = 0$. In that case the PPP deviation is expected to diverge, since the symmetry implied by $\mu_1 = -\mu_2$ and $\delta_{10} = \delta_{20}$ (see footnotes 12 and 13) ensures that the expected exchange rate is constant in the long run, while the expected PPP rate rises. Second, $\delta_{ppp} < 0$ also implies a diverging PPP deviation, as moving away from the PPP rate increases the probability of staying in that situation, so that the exchange rate is expected to get stuck in one regime after a while.

In the second part of the simulation experiment, we have to demonstrate that the limits of $E_{t-1}\{S_\tau - S_\tau^{ppp}|\tilde{r}_{t-1}\}$ and $V_{t-1}\{S_\tau - S_\tau^{ppp}|\tilde{r}_{t-1}\}$ do not depend on the initial condition I_{t-1}, that is, on S_{t-1} and $S_{t-1} - S_{t-1}^{ppp}$, as argued before. For that, we regress 100,000 simulated values of $S_\tau - S_\tau^{ppp}$ and $(S_\tau - S_\tau^{ppp})^2$ on randomly generated S_{t-1} and $S_{t-1} - S_{t-1}^{ppp}$ and their squares for various future times τ (both initial values are generated from the uniform distribution on (-0.5,0.5)). We find that for horizons up to about 1,000 the initial condition matters, but that for longer horizons it does not.[14] Hence, the limits of $E_{t-1}\{S_\tau - S_\tau^{ppp}|\tilde{r}_{t-1}\}$ and

---

[13] The reported results are based on $(\delta_{10}, \delta_{11}, \delta_{20}, \delta_{21}, \delta_{ppp}) = (4,0,4,0,0)$ and $(10,0,10,0,-1)$.

[14] The White (1980) heteroskedasticity robust F-tests for no effect of S_{t-1}, S_{t-1}^2, $S_{t-1} - S_{t-1}^{ppp}$ and $(S_{t-1} - S_{t-1}^{ppp})^2$ on $S_\tau - S_\tau^{ppp}$ and $(S_\tau - S_\tau^{ppp})^2$ for horizons 100, 500, 1000, 1500 and 2000 are as follows. For $S_\tau - S_\tau^{ppp}$ as dependent variable: $2{\cdot}10^4$ [p-value is 0.00], $9{\cdot}10^1$ [0.00], 1.08 [0.36], 0.64 [0.63], and 1.38 [0.24], respectively. For $(S_\tau - S_\tau^{ppp})^2$ as dependent variable: $7{\cdot}10^3$ [0.00], $4{\cdot}10^1$ [0.00], 2.46 [0.04], 0.68 [0.61], and 1.27 [0.28], respectively.
To verify that this gradual disappearance of the effect of the initial condition is not caused by misspecification of the linear regression model, we run a nonparametric regression (see Härdle and Linton (1994)) of $S_\tau - S_\tau^{ppp}$ on S_{t-1} and $S_{t-1} - S_{t-1}^{ppp}$ separately for the horizons just mentioned. The results, which are available from the author upon request, support our claim.

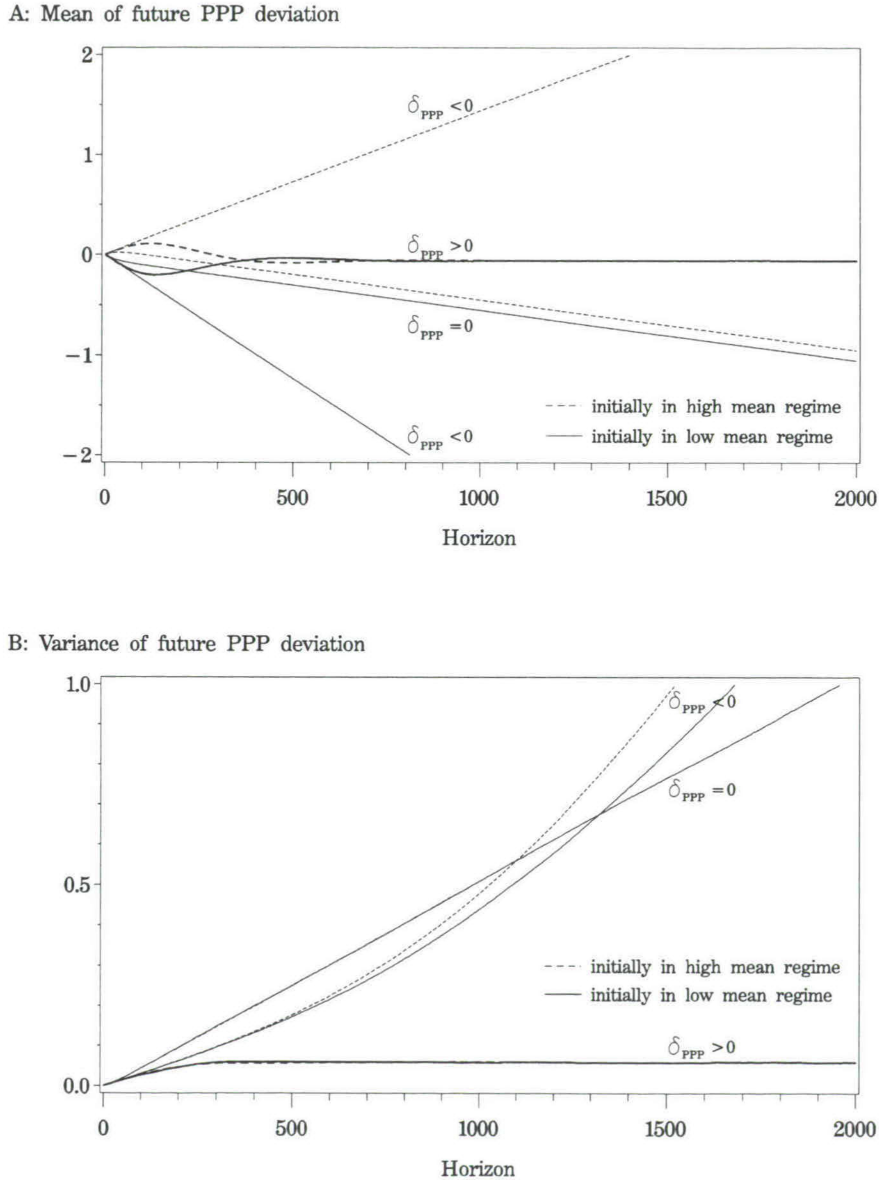


Figure 4.4: Behavior of future PPP deviations for different δ_{ppp} (measuring the strength with which the exchange rate is pulled towards PPP)

$V_{t-1}\{S_\tau - S_\tau^{ppp}|\tilde{r}_{t-1}\}$ indeed do not depend on I_{t-1}. Together with the conclusion from the first part of our simulation study, that both limits exist and do not depend on $\tilde{r}_{t-1}$, this shows that $\mu_1 < \mu_{ppp} < \mu_2$ and $\delta_{ppp} > 0$ indeed imply that PPP holds.

4.B P-values for PPP Tests

To decide whether the realized PPP tests $\widehat{\Pi}$ in subsection 4.3.2 are significant, we need the p-values. In this appendix we describe how we simulate them.

We generate 1,000 data sets, each containing one series of exchange rate levels S_t and one of PPP rate levels S_t^{ppp}. Both series are generated independently ($\delta_{ppp} = 0$), so that the data satisfy the null restriction of no PPP; a detailed description of the processes underlying both series follows below. For each data set, we estimate the regime-switching model with PPP, but without breaks in the duration of PPP deviations, thus (4.5) allowing for $\delta_{ppp} \neq 0$. This procedure yields 1,000 values for the PPP test statistic $\widehat{\Pi}$ in (4.6). The p-values for the three $\widehat{\Pi}$ that we estimate from the real data are the fractions of simulated $\widehat{\Pi}$ that exceed them.

We now discuss the 1,000 generated series of S_t and S_t^{ppp} in more detail. The length of both series is 1,217 time periods, the same as the length of the level series in the real data. Hence, our simulated p-values account for potential small-sample biases.

The process for S_t is the regime-switching process without PPP, as described in subsection 4.2.1. The true parameter values underlying each of the 1,000 series are the averages of the parameter estimates of the model without PPP for Germany, Japan and the U.K. (we will analyze the sensitivity of the p-values to this choice below).[15] The process starts from $S_0=0$.

Although the process for S_t^{ppp} is unspecified in section 4.2, we have to assume some process in this simulation exercise. As in appendix 4.A, we take a normal random walk with drift. Because under the null the PPP process is independent of the exchange rate process, it is not obvious how we should choose the values of the drift parameter μ_{ppp}. After all, if μ_{ppp} equals $E\{s_t\}$, then S_t and S_t^{ppp} seem to

[15]As can be verified from table 4.2, these averages are $\mu_1 = -0.29$, $\mu_2 = 0.14$, $\theta = 0.03$, $\sigma^2=2.51$, $\alpha=0.10$, $\beta=0.88$, $\nu^{-1}=0.19$, $p_{11}=0.983$ and $p_{22}=0.990$.

be related through their common trend, so that one will find many large values of simulated $\widehat{\Pi}$ and thus a large p-value for the realized $\widehat{\Pi}$, so that it becomes more difficult to reject the null. On the other hand, if μ_{ppp} is outside the interval (μ_1, μ_2), then the simulated $\widehat{\Pi}$ will very often be zero, leading to a low p-value for the realized $\widehat{\Pi}$ and to an easier rejection of the null. To get an objective value for μ_{ppp}, we use the International Financial Statistics (IFS) of the IMF. We first construct the PPP rates for all countries for which the IFS contains the consumer price index. For each of these 138 PPP rates, we then estimate μ_{ppp} and the corresponding variance σ^2_{ppp}. A randomly selected pair of these two estimates is taken as the true parameter pair underlying one of the 1,000 S_t^{ppp} series. The process starts from $S_0^{ppp}=0$.

A problem with this approach of generating S_t^{ppp} is that the OECD countries are relatively overrepresented in the IFS. Since OECD countries have quite stable PPP rates, this leads to too many μ_{ppp} close to zero. Hence, μ_{ppp} is too often close to $E\{s_t\}$, which is -0.01 for the true parameters for the S_t process.[16] As explained above, this similarity between μ_{ppp} and $E\{s_t\}$ makes the simulated p-values too high, so that it is more difficult to reject the null.

Having described the simulation procedure, we can now present the simulated p-values that we need in the main text. The column labeled "Basic case" in table 4.6 shows that they are 0.00 for Germany, 0.05 for Japan, and 0.01 for the U.K. Taking account of the fact that these simulated p-values likely overestimate the true ones, we conclude that the PPP test $\widehat{\Pi}$ is significant for all three exchange rates.

A potential problem with the p-values is that they are based on one specific set of true parameters for the exchange rate process and that the p-values are likely sensitive to that choice. First, if the parameters are changed such that $E\{s_t\}$ is more similar to μ_{ppp}, then the p-values will rise, as argued above. Second, if $\mu_2-\mu_1$ is made smaller, μ_{ppp} will more often be outside (μ_1, μ_2), thereby decreasing the p-values. In the remaining part of this appendix, we demonstrate that this sensitivity indeed exists, but that it is not problematic for our conclusion of rejecting the null.

To examine the sensitivity, we compute the p-values for several combinations

[16] See footnote 15, using that $E\{s_t\} = p_1\mu_1 + (1-p_1)\mu_2$, where the unconditional regime probability $p_1 = (1-p_{22})/(2-p_{11}-p_{22})$, as derived by Hamilton (1989).

Table 4.6: Simulated p-values for PPP tests and sensitivity to nuisance param.

	Basic case	Sensitivity analysis: Sensitivity to $E\{s_t\}$			Sensitivity to $\mu_2-\mu_1$		
Unconditional mean $E\{s_t\}$	-0.01	0.1	0	-0.1	0	0	0
Wedge regime means $\mu_2-\mu_1$	0.43	0.4	0.4	0.4	0.2	0.4	0.6
Critical value of PPP test $\widehat{\Pi}$	3.22	2.23	3.06	3.39	2.45	3.06	3.81
P-value Germany $(\widehat{\Pi}=5.46)$	[0.00]	[0.00]	[0.00]	[0.01]	[0.00]	[0.00]	[0.01]
Japan $(\widehat{\Pi}=3.14)$	[0.05]	[0.01]	[0.05]	[0.06]	[0.02]	[0.05]	[0.09]
U.K. $(\widehat{\Pi}=4.61)$	[0.01]	[0.01]	[0.01]	[0.03]	[0.00]	[0.01]	[0.04]

The column labeled "Basic case" contains the p-values that are used in the main text. These are computed from exchange rate and PPP rate processes simulated from parameter values that are equal to the average estimates of the model without PPP (see footnote 15).
The sensitivity analysis is based on different combinations of $E\{s_t\}$ and $\mu_2-\mu_1$. To transform each $(E\{s_t\},\mu_2-\mu_1)$ into the structural parameters μ_1 and μ_2, we assume for simplicity that the unconditional regime probabilities are both 0.5, so that $\mu_1 = E\{s_t\}-1/2(\mu_2-\mu_1)$ and $\mu_2 = E\{s_t\}+1/2(\mu_2-\mu_1)$. This is obtained by taking $p_{11}=p_{22}$, which we set at 0.987, the average of the values in footnote 15. The other parameters are kept at the average parameter estimates of the model without PPP (see footnote 15).
Further details about the simulation procedure are in Appendix 4.B.

of the nuisance parameters $E\{s_t\}$ and $\mu_2-\mu_1$, while using the same S_t^{ppp} series as before. The combinations are $E\{s_t\} = -0.1, 0, 0.1$ with $\mu_2-\mu_1$ held constant at 0.4 (to study the sensitivity regarding $E\{s_t\}$), and $\mu_2-\mu_1 = 0.2, 0.4, 0.6$ with $E\{s_t\}$ constant at 0 (to study the sensitivity with respect to $\mu_2-\mu_1$). These seem reasonable values given the estimates for the model without PPP in table 4.2, which imply that $(\widehat{E}\{s_t\},\widehat{\mu}_2-\widehat{\mu}_1)$ is $(0.01, 0.42)$ for Germany, $(-0.05, 0.43)$ for Japan, and $(0.01, 0.44)$ for the U.K.

Table 4.6 reports the p-values corresponding to each combination of nuisance parameters. It is clear that the p-values are indeed sensitive to both $E\{s_t\}$ and $\mu_2-\mu_1$. However, the results also show that this sensitivity is not problematic for our rejection of the null of no PPP. That is, even in the worst case the largest simulated p-value (for Japan) is quite small (0.09), particularly if one takes into account that the simulated p-values overestimate the true ones, as argued above.

4.C Estimation

We estimate the regime-switching model introduced in section 4.2 by maximum likelihood. In this appendix, we derive the likelihood function and show that it has a convenient recursive structure.

To obtain the likelihood function, we first need the density of the exchange rate change at time t conditional on only observable information. Let $p_{t-1}(s_t)$ denote this density evaluated at an exchange rate change equal to s_t.[17] It can be split up as

$$p_{t-1}(s_t) = \sum_{r_t, r_{t-1}=1,2} p_{t-1}(s_t \,|\, r_t, r_{t-1}) \cdot p_{t-1}(r_t, r_{t-1}). \tag{4.8}$$

We now discuss how to compute both terms on the right-hand-side.

The first term, $p_{t-1}(s_t \mid r_t, r_{t-1})$, denotes the density of the exchange rate change at time t evaluated at the value s_t conditional on I_{t-1} and on the current and previous regimes having values r_t and r_{t-1}, respectively. This t-density follows from formulas (4.2), (4.3) and (4.4). It is, however, not straightforward how to compute the conditional variance in (4.3), as this requires integrating out the regimes r_{t-1} and r_{t-2} in $\varepsilon_{t-1}^2 = \{s_{t-1} - [\mu_{r_{t-1}} + \theta(s_{t-2} - \mu_{r_{t-2}})]\}^2$. For that, we need $p_{t-1}(r_{t-1}, r_{t-2})$, the conditional probability that the two most recent regimes have values r_{t-1} and r_{t-2}. This probability is crucial, since all regime probabilities in the chapter can be derived from it. Using similar techniques as in Gray (1996a), we now show that this probability has a first-order recursive structure, which simplifies its computation a lot.

First, we write $p_{t-1}(r_{t-1}, r_{t-2})$ as

$$\begin{aligned} p_{t-1}(r_{t-1}, r_{t-2}) &= p_{t-2}(r_{t-1}, r_{t-2} \,|\, S_{t-1}^{ppp}, s_{t-1}) \\ &= p_{t-2}(r_{t-1}, r_{t-2} \,|\, s_{t-1}) \cdot \frac{p_{t-2}(S_{t-1}^{ppp} \,|\, r_{t-1}, r_{t-2}, s_{t-1})}{p_{t-2}(S_{t-1}^{ppp} \,|\, s_{t-1})}. \end{aligned} \tag{4.9}$$

This equation can be simplified by assuming that the ratio on the right-hand-side is one. That is, the information contained in the two recent exchange rate regimes is irrelevant for the distribution of S_{t-1}^{ppp} once all PPP levels through $t-2$

[17] We use the same symbol p_{t-1} for several densities (see (4.1) and (4.8)). The specific meaning of p_{t-1} is uniquely determined by the symbols we use in its argument. This results in a concise notation, which will prove useful in the remaining part of the chapter.

and all exchange rate levels through $t-1$ are known. This is reasonable, since the price levels underlying S_t^{ppp} are almost fixed in the short run. Given this assumption, we have

$$\begin{aligned} p_{t-1}(r_{t-1}, r_{t-2}) &= \frac{p_{t-2}(s_{t-1} \,|\, r_{t-1}, r_{t-2}) \cdot p_{t-2}(r_{t-1}, r_{t-2})}{p_{t-2}(s_{t-1})} \\ &= \frac{p_{t-2}(s_{t-1} \,|\, r_{t-1}, r_{t-2}) \cdot p_{t-2}(r_{t-1} | r_{t-2}) \cdot \sum_{r_{t-3}=1,2} p_{t-2}(r_{t-2}, r_{t-3})}{p_{t-2}(s_{t-1})}. \end{aligned} \tag{4.10}$$

Hence, the variables to compute $p_{t-1}(r_{t-1}, r_{t-2})$ are its previous values $p_{t-2}(r_{t-2}, r_{t-3})$ for $r_{t-3} = 1, 2$, the previous switching probability $p_{t-2}(r_{t-1}|r_{t-2})$ and the previous densities $p_{t-2}(s_{t-1} \mid r_{t-1}, r_{t-2})$ and $p_{t-2}(s_{t-1})$. This makes the computation of $p_{t-1}(r_{t-1}, r_{t-2})$ a first-order recursive process.

The second term on the right-hand-side of (4.8), $p_{t-1}(r_t, r_{t-1})$, is the conditional probability that the current and previous regimes have values r_t and r_{t-1}, respectively. It can be calculated from

$$p_{t-1}(r_t, r_{t-1}) = p_{t-1}(r_t \,|\, r_{t-1}) \cdot \sum_{r_{t-2}=1,2} p_{t-1}(r_{t-1}, r_{t-2}), \tag{4.11}$$

where $p_{t-1}(r_t \,|\, r_{t-1})$ follows from (4.7) and $p_{t-1}(r_{t-1}, r_{t-2})$ is given by (4.10).

Having discussed both terms on the right-hand-side of (4.8), we can now compute the density of interest, $p_{t-1}(s_t)$, being a mixture of four t-densities. This density can then be used to build the sample log-likelihood $\sum_{t=1}^{T} \log(p_{t-1}(s_t))$ with which all parameters in the regime-switching model can be estimated.

From a practical point of view, it is important to realize that the log-likelihood has a second-order recursive structure, similar to that of a standard one-regime AR(1)-GARCH(1,1) model. After all, for (4.11) one needs the current regime-switching probability $p_{t-1}(r_t \,|\, r_{t-1})$ and the first-order recursive probability $p_{t-1}(r_{t-1}, r_{t-2})$ for all eight combinations of (r_t, r_{t-1}, r_{t-2}); density (4.8) can then be computed from (4.11), the previous changes s_{t-1} and s_{t-2}, (4.10) and the previous variance $V_{t-2}\{\varepsilon_{t-1}\}$ in (4.3). This second-order recursiveness of $p_{t-1}(s_t)$ makes the calculation of the log-likelihood quite fast. To start up the recursive computation of the log-likelihood, we set the required variables equal to their expectation without conditioning on the information set.

4.D Regime Inference

As stated in footnote 10, the smoothed probability that the regime was r_t at time t, $p_T(r_t)$, can be computed recursively. More generally, any ex post regime probability $p_\tau(r_t)$, for a given future time $\tau \in \{t, t+1, \ldots, T\}$, can be calculated in a recursive manner. This claim, which we prove in this appendix, is based on the following recursive process for the two-regime ex post probability $p_\tau(r_t, r_{t-1})$ starting from the ex ante probability $p_{t-1}(r_t, r_{t-1})$.

Using an assumption similar to the one below (4.9), we can write $p_\tau(r_t, r_{t-1})$ for the four regime combinations as

$$\begin{aligned} p_\tau(r_t, r_{t-1}) &= p_{\tau-1}(r_t, r_{t-1} \,|\, s_\tau) \\ &= \frac{p_{\tau-1}(s_\tau | r_t, r_{t-1}) \cdot p_{\tau-1}(r_t, r_{t-1})}{\sum_{r_t, r_{t-1}=1,2} p_{\tau-1}(s_\tau | r_t, r_{t-1}) \cdot p_{\tau-1}(r_t, r_{t-1})}. \end{aligned} \tag{4.12}$$

Suppose first that $\tau = t$. Then $p_\tau(r_t, r_{t-1})$ follows directly from (4.12), as $p_{\tau-1}(r_t, r_{t-1})$ and $p_{\tau-1}(s_\tau | r_t, r_{t-1})$ are known from the estimation process (see appendix 4.C).

Hence, let us suppose from now on that $\tau > t$. The computation of (4.12) requires two inputs. The first one is the previous ex post probability $p_{\tau-1}(r_t, r_{t-1})$, which is known from the previous recursion for all combinations of r_t and r_{t-1}. The second ingredient of (4.12) is the density $p_{\tau-1}(s_\tau | r_t, r_{t-1})$ for all regime outcomes. Its computation requires a number of steps. We first write it as

$$p_{\tau-1}(s_\tau | r_t, r_{t-1}) = \sum_{r_\tau, r_{\tau-1}=1,2} p_{\tau-1}(s_\tau | r_\tau, r_{\tau-1}) \cdot p_{\tau-1}(r_\tau, r_{\tau-1} | r_t, r_{t-1}), \tag{4.13}$$

where we use that the conditional distribution of s_τ given $r_\tau, r_{\tau-1}$ does not depend on the earlier regimes r_t and r_{t-1}. This formula itself has two ingredients. The first one is the density $p_{\tau-1}(s_\tau | r_\tau, r_{\tau-1})$ for all regime combinations, which is known from the estimation process.

The second term needed in (4.13) is the $(\tau-t)$-period-ahead regime-switching probability $p_{\tau-1}(r_\tau, r_{\tau-1} | r_t, r_{t-1})$ for all regime combinations. Once it has been computed, it should be saved, since it will be needed in the next recursive step. Making use of the Markov structure of the regime process, it can be written in terms of $(\tau-1-t)$-period-ahead switching probabilities:

$$p_{\tau-1}(r_\tau, r_{\tau-1} | r_t, r_{t-1}) = \sum_{r_{\tau-1}, r_{\tau-2}=1,2} p_{\tau-1}(r_\tau, r_{\tau-1} | r_{\tau-1}, r_{\tau-2}) \cdot p_{\tau-1}(r_{\tau-1}, r_{\tau-2} | r_t, r_{t-1}). \tag{4.14}$$

Again, we have two ingredients. First, we need $p_{\tau-1}(r_\tau, r_{\tau-1}|r_{\tau-1}, r_{\tau-2})$ for all regime combinations. Due to the Markov property of the regime process, this switching probability does not depend on $r_{\tau-2}$. It equals

$$p_{\tau-1}(r_\tau, r_{\tau-1}|r_{\tau-1}, r_{\tau-2}) = p_{\tau-1}(r_\tau|r_{\tau-1}), \tag{4.15}$$

which is known from the estimation process.

The second ingredient of (4.14) is $p_{\tau-1}(r_{\tau-1}, r_{\tau-2}|r_t, r_{t-1})$ for all regime combinations. Using an assumption similar to the one below formula (4.9), we get

$$\begin{aligned} p_{\tau-1}(r_{\tau-1}, r_{\tau-2}|r_t, r_{t-1}) &= p_{\tau-2}(r_{\tau-1}, r_{\tau-2}|r_t, r_{t-1}, s_{\tau-1}) \\ &= \frac{p_{\tau-2}(s_{\tau-1}|r_{\tau-1}, r_{\tau-2}) \cdot p_{\tau-2}(r_{\tau-1}, r_{\tau-2}|r_t, r_{t-1})}{\sum_{r_{\tau-1}, r_{\tau-2}=1,2} p_{\tau-2}(s_{\tau-1}|r_{\tau-1}, r_{\tau-2}) \cdot p_{\tau-2}(r_{\tau-1}, r_{\tau-2}|r_t, r_{t-1})}, \end{aligned} \tag{4.16}$$

where we use that the conditional density of $s_{\tau-1}$ is independent of the previous regimes r_t, r_{t-1} once $r_{\tau-1}, r_{\tau-2}$ are given. We have two ingredients. First, the conditional density $p_{\tau-2}(s_{\tau-1}|r_{\tau-1}, r_{\tau-2})$ for all regime combinations. It is known from the estimation process. Second, we need the $(\tau-1-t)$-period-ahead switching probability $p_{\tau-2}(r_{\tau-1}, r_{\tau-2}|r_t, r_{t-1})$ for all regime combinations. This one was saved during the previous recursion, if $\tau > t+1$. If $\tau = t+1$, it equals one.

This completes the algorithm to compute (4.13), which is the second ingredient of (4.12). For each recursion one has to compute (4.16), use it together with (4.15) to compute (4.14) and use this to compute (4.13). Using this in (4.12) yields the ex post probability $p_\tau(r_t, r_{t-1})$ from $p_{\tau-1}(r_t, r_{t-1})$. Therefore, starting from the ex ante probability $p_{t-1}(r_t, r_{t-1})$ one can recursively compute the ex post probability $p_\tau(r_t, r_{t-1})$ and eventually the probability of interest $p_\tau(r_t)$.

4.E Forecasting

Subsection 4.3.5 deals with forecasting exchange rate levels S_τ at time $t-1$, where $\tau \geq t$. This appendix explains how to compute these forecasts.

As usual, we first decompose the exchange rate forecast as

$$E_{t-1}\{S_\tau\} = S_{t-1} + \sum_{i=t}^{\tau} E_{t-1}\{s_i\}. \tag{4.17}$$

To calculate $E_{t-1}\{s_i\}$, we rewrite s_i by repeated substitution of lags of (4.2) for the lagged changes. Since the innovations have zero expectation, this yields

$$E_{t-1}\{s_i\} = \sum_{r_i, r_{t-1}=1,2} p_{t-1}(r_i, r_{t-1}) \cdot \left(\mu_{r_i} + \theta^{i-(t-1)}(s_{t-1} - \mu_{r_{t-1}})\right), \tag{4.18}$$

where

$$p_{t-1}(r_i, r_{t-1}) = p_{t-1}(r_{t-1}) \cdot p_{t-1}(r_i \,|\, r_{t-1}), \tag{4.19}$$

where $p_{t-1}(r_{t-1})$ follows after summation of $p_{t-1}(r_{t-1}, r_{t-2})$ in (4.10) over r_{t-2}.

To compute the multi-period-ahead switching probability $p_{t-1}(r_i|r_{t-1})$ in (4.19), we first form the conditional one-period-ahead Markov transition matrices:

$$M_{t-1,j-1} = \begin{bmatrix} p_{t-1}(r_j{=}1 \,|\, r_{j-1}{=}1) & 1 - p_{t-1}(r_j{=}2 \,|\, r_{j-1}{=}2) \\ 1 - p_{t-1}(r_j{=}1 \,|\, r_{j-1}{=}1) & p_{t-1}(r_j{=}2 \,|\, r_{j-1}{=}2) \end{bmatrix}, \; j = t, \ldots, i. \tag{4.20}$$

For $j = t$, the elements of $M_{t-1,j-1}$ follow from (4.7); for $j > t$, we approximate $M_{t-1,j-1}$ by $M_{t-1,t-1}$. The theory of Markov processes for multi-period-ahead switching probabilities then implies that

$$p_{t-1}(r_i \,|\, r_{t-1}) = ({M_{t-1,t-1}}^{i-(t-1)})_{r_i r_{t-1}}. \tag{4.21}$$

Having explained how to calculate (4.19), we can now compute (4.18). Computation of (4.18) for all i and substitution in (4.17) then gives the forecast of interest $E_{t-1}\{S_\tau\}$.

Chapter 5

Have Exchange Rates Become More Closely Tied? Evidence from a New Multivariate GARCH Model

In this chapter we analyze the time-dependence of exchange rate correlations using a new multivariate GARCH model. This model consists of two parts. First, we transform the exchange rate changes into their principal components and specify univariate GARCH models for all components. Second, we use the inverse of the principal components construction to transform the conditional component moments back into those of the exchange rate changes themselves. The model is easy to estimate, as it requires only univariate GARCH estimations. Nevertheless, we find it outperforms the popular constant conditional correlations and factor GARCH models. We show that the major U.S. dollar exchange rates have become more loosely instead of closely tied since the eighties.

5.1 Introduction

Correlations are a key determinant of many financial decisions. For instance, investors in stocks need correlation assessments to reduce the riskiness of their portfolios, and correlations between exchange rates are important for internationally trading corporations and banks, as they have to hedge open foreign exchange positions. Several papers examine the correlations between stock returns, for instance, Bertero and Mayer (1990), Koch and Koch (1991), King, Sentana and Wadhwani (1994), Longin and Solnik (1995) and Darbar and Deb (1997). Surprisingly few papers, however, focus on exchange rate correlations. One notable example is Bollerslev (1990), who studies correlations between several European Monetary System (EMS) - U.S. dollar exchange rates. Therefore, in this chapter we also focus on exchange rate correlations.

Unlike Bollerslev (1990), however, we do not restrict the correlations to be constant. The reason is that exchange rate correlations are likely time-varying according to economic intuition. For example, suppose the U.K. joins the Exchange Rate Mechanism of the EMS. Then the correlation between the pound-dollar and, say, the mark-dollar exchange rates will rise. Secondly, a change in U.S. monetary policy such as the 1979 Volcker experiment also raises that correlation, since both the pound and the mark will change in the same way against the dollar. Therefore, we allow exchange rate correlations to vary over time. We find that the correlations between eight main U.S. dollar exchange rates have decreased since the eighties, so that exchange rates have become more loosely instead of closely tied.

When modeling high-frequency exchange rates, one has to take account of the well-known conditional heteroskedasticity in such data. The literature suggests various models for that, such as GARCH (see Bollerslev, Chou and Kroner (1992) for an overview), stochastic volatility (see Ghysels, Harvey and Renault (1996)), regime-switching GARCH (see Gray (1996a) and Chapter 2) and fractionally integrated GARCH (see Baillie, Bollerslev and Mikkelsen (1996)). For simplicity, we take GARCH, although the approach we will develop also works for any other volatility model.

Since we want to analyze correlations, a univariate GARCH model is insufficient, and a multivariate version is called for. In this chapter, we introduce a new multivariate GARCH model that is more suitable for a detailed correlation analysis than existing multivariate GARCH variants, as we will explain below. The basic idea of our model stems from the fact that it is the correlations between exchange rates that make multivariate GARCH modeling more difficult than univariate GARCH. Therefore, in the first step of our approach, we remove all unconditional correlations by taking principal components of the exchange rate changes. The conditional mean and variance of each principal component are specified by a univariate GARCH model. In the second step, the inverse of the principal components construction is used to transform the conditional moments of the principal components into the conditional mean and variance of the exchange rate changes themselves. Since this step requires no further estimation, our indirect approach makes multivariate GARCH estimation as easy as several univariate GARCH estimations.

The remaining part of this introduction presents a brief overview of the literature on multivariate GARCH and explains the contribution of this chapter in more detail.

In the GARCH literature so far, extending univariate to multivariate GARCH has been a main endeavour. The reason is that one has to model not only conditional variances, but also all conditional covariances. This can easily lead to an enormous number of parameters. Hence, multivariate GARCH modeling amounts to finding a parsimonious specification of the conditional covariance matrix that does not imply an unacceptable loss of generality.

In this respect, the diagonal model of Bollerslev, Engle and Wooldridge (1988) and the BEKK model of Engle and Kroner (1995) are useful for low-variate systems. However, estimation becomes difficult for higher-variate systems. For instance, in our eight-variate empirical application, one would have to estimate more than a hundred parameters. From a computational point of view, our model is more convenient, as it requires only univariate GARCH estimations.

Another computationally attractive model is the popular Bollerslev (1990) constant conditional correlations model. For our study, however, the model is not suitable, as we want to focus on the dynamics in exchange rate correlations. As indicated above, economic intuition shows that such dynamics are very likely present. This is clearly supported by our data. In this sense, our model is preferable, as it can explain time variation in correlations, leading to a better fit.

A fourth class of existing multivariate GARCH models is factor GARCH; see Diebold and Nerlove (1989), Engle, Ng and Rothschild (1990), Ng, Engle and Rothschild (1992), King, Sentana and Wadhwani (1994) and Fiorentini, Sentana and Shephard (1998). Two reasons behind the success of factor GARCH are that such models are computationally tractable and that, in contrast to the Bollerslev (1990) model, they can capture some time-variation in the conditional correlations. However, the fit of conditional variances and correlations is not as good as that of the model we propose. The explanation will become clear in the next paragraph.

Although our model has practical advantages over existing models, it also has a sound theoretical basis. This stems from the fact that the model is a factor GARCH model, although not in its traditional form as used above. Usual factor GARCH models are based on the theory that only a few unobserved variables, the factors, govern all exchange rates. Our model, on the other hand, uses as many factors as exchange rates. We demonstrate that this is the reason for the empirical outperformance of our model with respect to usual factor GARCH.

Taking the maximum number of factors in factor GARCH has two advantages. First, this choice turns out to be significantly optimal. Hence, our model solves a major problem of factor GARCH, namely the choice of the number of factors.

The second advantage concerns estimation. To estimate usual factor GARCH models, one commonly takes a two-step estimation method to avoid a complicated simultaneous procedure (see Engle et al. (1990), among others). Correction of the second-step standard errors for first-step estimation inaccuracy, however, is difficult and thus often ignored, leading to biased inference. Since there is no estimation in the second step of our method, we do not have this potentially serious problem.

Our model yields the following conclusions regarding the development of exchange rate correlations over time. First, we find that correlations between the main U.S. dollar exchange rates were decreasing in the years after the first oil shock, were increasing at the end of the seventies, and that they were highest in the eighties.

Second, concerning the central question of the chapter, we show that exchange rates have become more loosely instead of closely tied since the eighties. This is caused by the 1992 collapse of the Exchange Rate Mechanism of the EMS, which made several European exchange rates less correlated. Moreover, the EMS - yen correlations have become lower because of the coexistence of more stable EMS - U.S. dollar rates and a long swing in the yen - dollar rate in the nineties.

The plan for the rest of the chapter is as follows. In the next section, we introduce our multivariate GARCH model. Section 5.3 explains why that model is a special factor GARCH model with the maximum number of factors. In section 5.4 we present the empirical results and analyze the time-variation in the correlations. Section 5.5 concludes.

5.2 A New Multivariate GARCH Model

In the first subsection, we develop the new multivariate GARCH model to be used for our study on exchange rate correlations. In subsection 5.2.2 we explain how to estimate the model. In the final subsection, we examine the implications of our model for the conditional correlations.

5.2.1 The Model

The basic idea behind the model is based on the fact that it is the correlations between exchange rates that make multivariate GARCH models more complex than univariate ones. Therefore, we first remove the (unconditional) correlations by transforming the exchange rate changes into their principal components. We bring in the GARCH effects through these components instead of directly through the exchange rate changes themselves. In the second step, we then transform the principal component moments into the moments of the exchange rate changes, which we are interested in.

To describe the model, we need the following notation. Let S_t denote the vector of logarithms of I (nominal) spot exchange rates at time t, where each exchange rate is defined as the domestic currency price of one unit of foreign currency. We concentrate on the I-vector s_t consisting of the (percentage) exchange rate changes $s_{it} = 100(S_{it} - S_{it-1})$. Thus, s_{it} is the depreciation of the domestic currency with respect to currency i.[1] All exchange rate changes up to and including time $t-1$ form the information set I_{t-1}. Finally, we assume that s_t is conditionally normally distributed. Therefore, we only concentrate on its conditional mean and variance.

In the first part of our model, we concentrate on the I-vector of principal components defined by

$$f_t = W' s_t, \tag{5.1}$$

where the weighting matrix W is the unique (apart from column exchanges) orthogonal $I{\times}I$ eigenvector matrix of the unconditional variance $V\{s_t\}$. This

[1] This notation is closely related to the notation in Chapters 2 to 4. The only difference is that now S_t and s_t denote vectors, whereas in the three previous chapters they were scalars.

transforms the correlated exchange rate changes into their (unconditionally) uncorrelated principal components.

To specify $E_{t-1}\{f_t\}$ and $V_{t-1}\{f_t\}$, the mean and variance of f_t conditional on the information set I_{t-1}, we use a standard, univariate AR(1)-GARCH(1,1) model for each principal component separately. We complete the matrix $V_{t-1}\{f_t\}$ by assuming that the off-diagonal elements are zero; this assumption is quite common in the literature (see Engle et al. (1990) and Ng et al. (1992), among others). In summary, we specify the conditional moments of f_t by

$$\begin{aligned} E_{t-1}\{f_{kt}\} &= \mu_k + \theta_k(f_{kt-1} - \mu_k) \\ V_{t-1}\{f_{kt}\} &= \omega_k + \alpha_k(f_{kt-1} - E_{t-2}\{f_{kt-1}\})^2 + \beta_k V_{t-2}\{f_{kt-1}\} \\ Cov_{t-1}\{f_{kt}, f_{lt}\} &= 0, \end{aligned} \tag{5.2}$$

for principal components $k, l = 1, \ldots, I$, $k \neq l$. This completes the first part of the model.

In the second part of the model, we have to transform the conditional moments of the principal components into the ones for the exchange rate changes themselves, as it is the exchange rates that we are mainly interested in. This transformation is straightforward, as (5.1) and the orthogonality of the weighting matrix W imply that

$$\begin{aligned} E_{t-1}\{s_t\} &= W E_{t-1}\{f_t\} \\ V_{t-1}\{s_t\} &= W V_{t-1}\{f_t\} W'. \end{aligned} \tag{5.3}$$

This completes the second and final part of our multivariate GARCH model. Hence, the complete multivariate GARCH model is given by (5.1), (5.2) and (5.3).

5.2.2 Estimation

In this subsection we describe how to estimate our model. The first part of the model, represented by (5.1) and (5.2), can be estimated by principal components analysis on the sample covariance matrix of s_t, followed by maximum likelihood estimation of the normal univariate GARCH models for each sample principal component separately. Remarkably, this is all that is needed to estimate the model; the second part of the model, the inverse transformation (5.3), requires no further estimation, as the weighting matrix W has already been estimated in the first step. Hence, estimation of the full multivariate GARCH system is essentially as simple as several univariate GARCH estimations. This makes our model attractive from a practical point of view, as several other multivariate GARCH models, such as the diagonal and BEKK models mentioned in the introduction, are more difficult to estimate.

5.2.3 Implications for the Conditional Correlations

The focus of the chapter is the development of exchange rate correlations over time. In the introduction we have argued that our model improves over the Bollerslev (1990) constant conditional correlations model in this respect, because our model allows for time-variation in the conditional correlations. However, our model also imposes some structure on the correlations. In this subsection, we examine whether this structure is reasonable.

In our model, the time-variation in the conditional correlations is completely driven by the time-variation in the conditional variances of the principal components. This follows directly from the conditional variance formula in (5.3) and the diagonality of $V_{t-1}\{f_t\}$.

To see whether such a structure is reasonable, consider the following stylized example. Suppose we have $I = 2$ U.S. dollar exchange rate changes, namely the U.K. pound (s_{1t}) and the German Mark (s_{2t}). Assume that both have unit unconditional variance. This implies that the principal components are

$$\begin{aligned} f_{1t} &= \sqrt{1/2} \cdot s_{1t} + \sqrt{1/2} \cdot s_{2t} \\ f_{2t} &= \sqrt{1/2} \cdot s_{1t} - \sqrt{1/2} \cdot s_{2t} = \sqrt{2} \cdot s_{1t} - f_{1t}, \end{aligned} \tag{5.4}$$

where the joint component f_{1t} represents the EMS - U.S. dollar exchange rate and the difference component f_{2t} represents the deviation of the U.K. pound from the EMS. Using the variance formula in (5.3), straightforward calculations show that the conditional correlation between the U.K. pound and the German mark exchange rate changes equals

$$\rho_{t-1}\{s_{1t}, s_{2t}\} = \frac{\frac{1}{2}V_{t-1}\{f_{1t}\} - \frac{1}{2}V_{t-1}\{f_{2t}\}}{\frac{1}{2}V_{t-1}\{f_{1t}\} + \frac{1}{2}V_{t-1}\{f_{2t}\}}. \tag{5.5}$$

To analyze whether this specification is reasonable, we analyze the effects of two different policy changes. First, suppose the U.K. joins the Exchange Rate

Mechanism (ERM) of the EMS. Then the U.K.-EMS component f_{2t} becomes more stable, so that $V_{t-1}\{f_{2t}\}$ falls and the correlation $\rho_{t-1}\{s_{1t}, s_{2t}\}$ rises, as expected.

The second policy change we consider is a change in U.S. monetary policy, which increases the conditional variance of the U.S. dollar versus both EMS currencies. According to the model, the increase in $V_{t-1}\{f_{1t}\}$ raises the intra-EMS correlation $\rho_{t-1}\{s_{1t}, s_{2t}\}$. This is realistic, as both the pound and the mark change in the same way against the dollar after the policy shift.

Although we admit that the previous example is simple, it does show that the restrictions our model imposes on the conditional correlations are quite reasonable. In this sense, our model is preferable over the popular Bollerslev (1990) model. After all, that model restricts the conditional correlations to stay constant, even after important policy changes such as the ones discussed above.

5.3 Relation with Factor GARCH

In the previous section, we have seen that our model has some advantages over three existing multivariate GARCH models, namely the diagonal, the BEKK and the constant conditional correlations model. In this section we relate the model to the fourth class of existing models, namely factor GARCH. It turns out that our model is a factor GARCH model, albeit not in its traditional form. The usual factor GARCH model assumes that there are only a few factors that govern all exchange rates. In contrast, our model takes as many factors as there are exchange rates. This claim is proved in subsection 5.3.1.

Although our model uses many more factors than usual factor GARCH, this does not necessarily mean that our model is substantially better. Maybe the inclusion of extra factors does not lead to a much better fit and only complicates the model. In subsection 5.3.2 we argue that this is not the case.

5.3.1 A Special Factor GARCH Model

In this subsection we demonstrate that our model of subsection 5.2.1 is a factor GARCH model with as many factors as exchange rates. We only address the main points of this derivation; the complete derivation is in the appendix.

The central idea of a K-factor GARCH model is that there are K underlying variables, the factors, that govern all I exchange rate changes. More formally, the exchange rate innovation $\varepsilon_t = s_t - E_{t-1}\{s_t\}$ has a systematic and an unsystematic part, where the systematic part is a linear combination of K unobserved factors φ_{kt}:

$$\varepsilon_t = \Lambda\varphi_t + \upsilon_t, \tag{5.6}$$

where $\varphi_t = (\varphi_{1t}, \ldots, \varphi_{Kt})'$ is the K-vector of common factors with a time-varying conditional covariance matrix, Λ is the $I \times K$ full-column-rank matrix of factor loadings, and υ_t denotes the vector of unsystematic, exchange rate specific changes with a covariance matrix that is constant over time.

There are two problems with a direct implementation of the factor idea. The first problem is that the systematic and unsystematic innovations, φ_t and υ_t, are not observed separately, so that Λ is, in general, not directly estimable. As shown by Engel et al. (1990), this problem can be solved by substituting the vector of unobserved factors φ_t by an expression based on an observed K-vector that is closely related to the factors, in the sense that the conditional variance of the k-th component of this factor representing vector is perfectly correlated with that of the k-th factor φ_{kt} (see the end of footnote 11 for a formalization of this). Similar to the existing literature (see Ng et al. (1992), among others), we take K principal components of s_t to form this factor representing vector, and we assume that they are conditionally uncorrelated and that each of them follows a normal AR(1)-GARCH(1,1) model. Hence, the factor representing vector is a K-dimensional subvector of f_t, the vector of all I principal components defined by (5.1) and modeled by (5.2). For simplicity of notation, let us denote this subvector of f_t also by f_t, and let W also denote the $I{\times}K$ full-column-rank matrix of component weights that defines the subvector by $f_t = W's_t$.

The second problem with a direct implementation of the factor idea is caused by a rotational indeterminacy in the factors φ_t in (5.6); this makes the matrix of factor loadings Λ unidentified. The appendix shows that this problem is also present after the move to the factor representing vector. To solve the problem, we normalize Λ by $W'\Lambda = I_K$, where I_K is the $K{\times}K$ identity matrix. This normalization will appear to be crucial for proving the claim that our model is a factor GARCH model with $K = I$ factors.

Having solved both problems, we can derive the commonly-used K-factor

GARCH formulas for the two conditional moments of interest:

$$\begin{aligned} E_{t-1}\{s_t\} &= \gamma + \Lambda E_{t-1}\{f_t\} \\ V_{t-1}\{s_t\} &= \Omega + \Lambda V_{t-1}\{f_t\}\Lambda', \end{aligned} \tag{5.7}$$

where γ and Ω are time-constant parts in the mean and variance, respectively. These moment specifications hold for all $K \in \{1, \dots, I\}$. Note, however, that for $K = I$ the constants γ and Ω are zero. After all, in that case $\Lambda f_t = \Lambda W' s_t = s_t$, because the normalization $W'\Lambda = I_K$ then implies that $\Lambda = (W')^{-1}$.

Although some similarities with our model of section 5.2 have already become clear, it may not yet be clear that our model exactly equals the I-factor GARCH model. The final link is provided by our factor GARCH normalization $W'\Lambda = I_K$ and the orthogonality of W. They imply that $\Lambda = (W')^{-1} = W$. Hence, relation (5.7), where γ and Ω are zero due to $K = I$, is the same as the second part of our model given by (5.3). Because the models for the I principal components are also the same, our model is indeed a special factor GARCH model in which the number of factors equals the number of exchange rates.

5.3.2 Advantages over the Usual Factor GARCH Model

From the previous subsection we know that our model uses many more factors than usual factor GARCH. In this subsection, we demonstrate that including these extra factors is useful by showing that our model overcomes two important problems with the empirical implementation of usual factor GARCH models. These problems are the choice of the number of factors and the difficult correction of standard errors in the two-step method that is commonly used to estimate factor GARCH.

The first problem is the choice of the number of factors K, or, equivalently, the number of principal components. This problem originates from a trade-off between generality and simplicity. On the one hand, increasing K leads to a more general model, but, on the other hand, it makes the model more complicated.

To alleviate this problem, there are several ad hoc criteria for selecting K (see Bartholomew (1987)). The most popular one is the Kaiser-Guttman rule, which states that one should select only those principal components that have a larger variance than the average variance of the exchange rate changes. As all other rules, this one yields very few components. For instance, in our eight-variate empirical application, it would select only one.

To investigate whether the neglect of components is serious, we estimate the factor GARCH model for all possible K, using the exchange rate data that we will describe in subsection 5.4.1. The results, which are described in detail in subsection 5.4.5, show that using less than I components is strongly rejected. Some components turn out to be essential for a good description of the conditional variances, while other principal components, which do not improve the variance fit much, turn out to be important for the correlation fit. The usual factor GARCH model neglects many of these important components. This demonstrates the dangers involved in the popular rules for choosing the number of factors. According to our results, the correct rule is to use as many factors as possible. Since our model does exactly that and is, nevertheless, easy to estimate, it solves the problem of choosing K in usual factor GARCH.

The second problem with the empirical implementation of the usual factor GARCH model is the difficult correction of standard errors in the two-step method that is commonly used for estimation. To clarify this, we first describe this two-step method. The first step is similar to the first step of our method as described in subsection 5.2.2. The only difference is the number of univariate GARCH models for the principal components that one has to estimate: I in our model and K for a K-factor GARCH model, as follows from the previous subsection.

The second step in the estimation of usual factor GARCH, however, is essentially different. After substitution of the first step estimates for $E_{t-1}\{f_t\}$ and $V_{t-1}\{f_t\}$ in (5.7), the usual factor GARCH model requires estimation of the parameters γ, Ω and Λ to obtain estimates for the moments of interest, $E_{t-1}\{s_t\}$ and $V_{t-1}\{s_t\}$.[2] Because one uses only estimates instead of the true values of $E_{t-1}\{f_t\}$

[2]Most researchers use univariate techniques for this second estimation step. That is, for each exchange rate i, they use maximum likelihood based on conditional normality of s_t with mean and variance implied by the corresponding elements of (5.7). As Ng et al. (1992) admit, such univariate estimation sacrifices efficiency. The reason for not doing multivariate maximum likelihood is that this would lead to too many parameters to be estimated at once. After all, γ, Ω and Λ have $I + I^2 + I \cdot K$ unknown elements. This may indeed be too much, if one does not take account of all restrictions that the factor GARCH model puts on γ, Ω and Λ. These restrictions are our normalization $W'\Lambda = I_K$, which also implies $W'\gamma = 0$, and the definition of Ω (see below (5.14), with the additional assumption of a diagonal $V\{v_t\}$ that we use in the empirical section). They greatly reduce the number of free parameters. For instance, for $K=7$ and $I=8$, they lead to 16 free parameters instead of 128! Therefore, multivariate estimation is not that difficult, and we prefer it over the univariate techniques used in the literature so far.

and $V_{t-1}\{f_t\}$, the second step standard errors have to be corrected for the first step estimation inaccuracy. This is complicated, as Lin (1992) shows. Therefore, many authors do not correct them and use the potentially seriously biased second step standard errors for inference. In this respect, our model is preferable, because the second step is a linear transformation without any estimation (see (5.3)). Hence, our model involves neither difficult standard error correction, nor the use of biased standard errors.

In summary, the fact that our model employs many more factors than usual factor GARCH is very useful. First, by using the optimal number of factors, the model yields a better fit. Second, estimation is easier than for any other factor GARCH model, as our approach does not require a second estimation step.

5.4 Empirical Results

In this section we use our multivariate GARCH model to analyze the development of exchange rate correlations over time. First, we describe the data and motivate the choice for our model empirically. Then we estimate the model. In subsection 5.4.3 we addresses the central question of the chapter, namely whether exchange rates have become more closely tied. Then we check whether the model captures the main characteristics of the data and in subsection 5.4.5 we compare the fit of our model with some benchmark models, namely the Bollerslev (1990) model with constant conditional correlations and factor GARCH models for all possible numbers of factors.

5.4.1 Data

We use U.S. dollar exchange rates of $I=8$ currencies, namely, the Belgian franc, Canadian dollar, French franc, German mark, Italian lira, Japanese yen, Dutch guilder and the British pound. These include all major exchange rates. Moreover, some of them are highly correlated (the EMS rates), while others are much less correlated; this variety allows us to get a fairly complete picture of the behavior of the conditional correlations. We have 1,216 weekly observations for the weekly changes s_t from April 1974 to July 1997. All rates have been obtained from Datastream.

In table 5.1 we report some descriptive statistics; the notes below the table contain the definitions. The substantial cross-currency correlations in the first panel motivate the use of a multivariate model instead of univariate ones.

In the second panel of table 5.1, we test for autocorrelation in the exchange rate changes. We find significantly positive first-order autocorrelation in the core EMS exchange rate changes (we always use a significance level of 5%).[3] For this reason, we have allowed for a first-order autoregressive term in (5.2), the model for the principal components. Estimates for higher-order autocorrelations are not reported separately, but are combined in Box-Pierce type statistics $\widetilde{Q}_{10}$; they indicate that higher-order autoregressive terms are unnecessary.

The third panel of table 5.1 deals with the dynamics of the second moments. The first two rows contain measures for the time-variation in the squared exchange rate changes. Both measures point at conditional heteroskedasticity. The next two rows of panel three contain similar autocorrelation measures, but now regarding the cross products instead of the squares. Since there are seven cross products for each exchange rate series, we have taken the average to save space. The results show clear evidence of time-variation in the conditional covariances. Hence, the data motivate the use of a multivariate GARCH model.

A popular multivariate GARCH model is the Bollerslev (1990) model, which assumes that all conditional correlations for the exchange rate changes are constant over time. In the last row of table 5.1, we test this restriction as follows. First, we estimate a univariate GARCH model for each series of exchange rate changes and construct conditional correlation estimates by taking the product of the normalized residuals. Then we regress the estimated conditional correlations for time t on a the vector $(1, t, t^2)'$ and test whether the two slope parameters are zero (see the notes below table 5.1 for further details). The results show that there is clear time-dependence in the conditional correlations. This is not surprising. First, economic intuition tells us that correlations may well be time-varying (see the policy examples in subsection 5.2.3). Second, Bollerslev (1990) already shows

[3]Our evidence of first-order autocorrelation is in contrast with conclusions of many earlier studies. For instance, West and Cho (1995) conclude from heteroskedasticity corrected Ljung-Box statistics of orders 10, 50 and 90 that five major U.S. dollar exchange rate changes are serially uncorrelated, with the possible exception of the yen. Indeed, if we had only used the aggregate Box-Pierce type measure $\widetilde{Q}_{10}$ in table 5.1, we would have concluded the same, thereby overlooking the significant first-order autocorrelation in all core EMS exchange rate changes. Hence, our additional check for only first-order autocorrelation is useful.

Table 5.1: Moments of exchange rate changes and autocorrelation tests

	Bel	Can	Fra	Ger	Ita	Jap	Neth	U.K.
Mean	0.00	-0.03	-0.02	0.03	-0.08	0.07	0.02	-0.03
Variance	2.11	0.38	2.06	2.14	2.06	2.11	2.09	2.13
Skewness	-0.26	-0.57	-0.25	-0.14	-0.59	0.53	-0.14	-0.40
Excess kurtosis	1.92	6.78	2.34	1.70	6.38	2.01	1.90	3.00
Cross-currency corr. $\overline{\rho}$	0.71	0.13	0.71	0.72	0.64	0.47	0.73	0.59
Autocorr. s_{it}: ρ_1	0.07* (0.03)	0.01 (0.04)	0.06* (0.03)	0.07* (0.03)	0.02 (0.04)	0.05 (0.04)	0.07* (0.03)	0.04 (0.04)
Autocorr. s_{it}: $\widetilde{Q}_{10}$	14.82 [0.14]	7.35 [0.69]	14.63 [0.15]	14.07 [0.17]	11.78 [0.30]	22.57* [0.01]	13.31 [0.21]	6.05 [0.81]
Autocorr. s_{it}^2: ρ_1^s	0.09* (0.03)	0.15* (0.03)	0.05 (0.03)	0.04 (0.03)	0.10* (0.03)	0.20* (0.03)	0.07* (0.03)	0.21* (0.03)
Autocorr. s_{it}^2: Q_{10}^s	48.66* [0.00]	36.53* [0.00]	52.49* [0.00]	57.60* [0.00]	134.20* [0.00]	92.03* [0.00]	56.24* [0.00]	151.82* [0.00]
Autocorr. $s_{it} \cdot s_{jt}$: $\overline{\rho_1^c}$	0.07* (0.03)	0.07* (0.03)	0.07* (0.03)	0.05 (0.03)	0.08* (0.03)	0.05 (0.03)	0.06* (0.03)	0.06* (0.03)
Autocorr. $s_{it} \cdot s_{jt}$: $\overline{Q_{10}^c}$	57.01* [0.00]	18.46* [0.05]	61.79* [0.00]	57.76* [0.00]	61.95* [0.00]	34.64* [0.00]	55.81* [0.00]	82.21* [0.00]
Constancy of conditional corr.	19.76* [0.00]	21.70* [0.00]	22.34* [0.00]	20.30* [0.00]	21.38* [0.00]	17.94* [0.00]	20.38* [0.00]	22.61* [0.00]

Standard errors in parentheses and p-values in square brackets; * is significant at 5% level.
The correlation measure $\overline{\rho}$ is the average of the sample correlation coefficients of the series under consideration with all seven other series.
The first-order autocorrelation, ρ_1, is estimated as the slope coefficient in a regression of the change in exchange rate i, s_{it}, on the first lagged change, s_{it-1}, and a constant. The standard errors are based on White's (1980) heteroskedasticity-consistent asymptotic covariance matrix.
$\widetilde{Q}_{10}$ denotes a modified Box-Pierce type statistic that combines the first ten autocorrelations. Following Pagan and Schwert (1990), it is defined as the sum of the first ten squared normalized autocorrelation estimates, where the normalizing factors are the heteroskedasticity-consistent standard errors of the autocorrelation estimates. $\widetilde{Q}_{10}$ is asymptotically χ^2_{10} distributed.
The first-order autocorrelation in the squared changes, ρ_1^s, and the Box-Pierce type statistic for the squared changes, Q_{10}^s, are similarly defined as ρ_1 and $\widetilde{Q}_{10}$, respectively, although without the heteroskedasticity correction.
The seven first-order autocorrelations of the cross products $s_{it} \cdot s_{jt}$ $(j \neq i)$ are averaged to save space; this average is denoted by $\overline{\rho_1^c}$. The number in parentheses is also the average standard error. Similarly, $\overline{Q_{10}^c}$ denotes the mean of the seven Box-Pierce type statistics of the cross products; its p-value is based on a χ^2_{10} distribution.
We test for constancy of the conditional correlations $\rho_{t-1}\{s_{it}, s_{jt}\}$ by testing the constancy of $Cov_{t-1}\{\varepsilon_{it}, \varepsilon_{jt}\}/(V_{t-1}\{\varepsilon_{it}\}V_{t-1}\{\varepsilon_{jt}\})^{1/2}$, where ε_{it} is the innovation in a univariate normal-AR(1)-GARCH(1,1) model for s_{it}. The test amounts to regressing the estimated correlation, $\widehat{\varepsilon}_{it}\widehat{\varepsilon}_{jt}/(\widehat{V}_{t-1}\{\varepsilon_{it}\}\widehat{V}_{t-1}\{\varepsilon_{jt}\})^{1/2}$, on a constant, t and t^2, and then computing the Wald statistic for no effect of t and t^2. For space considerations, we only report the average over the seven possible Wald statistics for each i. The critical values are based on a χ^2_2 distribution.

Table 5.2: Principal component weights

	EMS	Jap	U.K. – EMS	Ita – EMS	Can	Fra – EMS	Bel – G+N	Neth – Ger
Belgium	0.41	-0.09	-0.23	-0.22	0.03	-0.40	0.75	0.03
Canada	0.03	-0.05	0.12	0.10	0.99	-0.02	-0.01	-0.01
France	0.40	-0.07	-0.15	-0.05	0.03	0.88	0.19	-0.03
Germany	0.42	-0.06	-0.23	-0.23	0.02	-0.18	-0.44	-0.70
Italy	0.36	-0.19	-0.02	0.89	-0.11	-0.13	-0.02	-0.01
Japan	0.28	0.94	0.14	0.09	0.01	-0.03	0.02	0.01
Netherlands	0.41	-0.07	-0.22	-0.20	0.03	-0.12	-0.46	0.71
U.K.	0.34	-0.22	0.89	-0.19	-0.11	-0.03	-0.00	-0.00
Variance	11.60	1.31	0.85	0.58	0.36	0.22	0.13	0.04
Expl. variance	76.87	8.70	5.65	3.83	2.41	1.43	0.83	0.27

Each column contains the weights of the individual exchange rates changes in the sample principal components. The eight weighting vectors, named according to the dominating currencies, form the weighting matrix W in (5.1). Hence, W is the matrix of eigenvectors of the sample covariance matrix of the exchange rate changes (normalized at length one, so that the "weights" do not sum to one).
"Variance" denotes the sample variance of a principal component, which is equal to the corresponding eigenvalue.
"Expl. variance" denotes the percentage of the total variance explained by a principal component, that is, the sample variance of the component divided by the sum of the sample variances of the individual exchange rate changes (called the "total variance").

that conditional correlations differ between the pre-EMS and the EMS period. In addition, Andersen, Bollerslev, Diebold and Labys (1999) also find strong empirical evidence of time-varying conditional correlations in exchange rates. This motivates why we use our model instead of the Bollerslev (1990) model, since our model can capture time-variation in the conditional correlations.

5.4.2 Estimation Results

In this subsection we estimate our multivariate GARCH model. As the second part of this model involves no estimation (see subsection 5.2.2), we only concentrate on the first part, that is, the principal components construction and the univariate GARCH estimations for each component.

To construct the principal components vector f_t, we define the weighting matrix W in (5.1) by the matrix of eigenvectors of the sample covariance matrix of s_t. The upper panel of table 5.2 presents the columns of W, which are the weighting

Table 5.3: Estimation results for the principal components

		EMS	Jap	U.K. – EMS	Ita – EMS	Can	Fra – EMS	Bel – G+N	Neth – Ger
Mean	μ	-0.01	0.10*	-0.03	-0.04*	-0.01	-0.02*	-0.00	-0.00
		(0.08)	(0.03)	(0.03)	(0.02)	(0.02)	(0.01)	(0.00)	(0.00)
Autocorr.	θ	0.06	0.07*	0.09*	-0.00	0.02	-0.15*	-0.30*	-0.25*
		(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)	(0.03)
Cond. var. intercept	ω	0.15	0.08*	0.16*	0.09*	0.04*	0.00*	0.00*	0.00
		(0.08)	(0.03)	(0.05)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)
ARCH	α	0.16*	0.13*	0.10*	0.38*	0.16*	0.33*	0.26*	0.02*
		(0.03)	(0.03)	(0.03)	(0.06)	(0.03)	(0.05)	(0.03)	(0.00)
GARCH	β	0.85*	0.82*	0.72*	0.51*	0.74*	0.80*	0.81*	0.98*
		(0.03)	(0.05)	(0.08)	(0.05)	(0.05)	(0.02)	(0.01)	(0.00)
Log-likelihood		-3131	-1836	-1597	-1128	-1070	-596	-112	652

Standard errors in parentheses; * is significant at 5% level.
The estimated model is (5.2) without the conditional covariance equation. To start-up the conditional variance, we use a separate parameter, which is not reported. Standard errors are not corrected for the fact that we use only an estimate of the weighting matrix W, because our focus is on the conditional moments of the exchange rate changes, not the intermediate GARCH estimation results for the principal components.

vectors for the principal components. Each of the eight components has a name that indicates the dominating currencies in it. Hence, the components are called EMS, Jap, U.K.-EMS, Ita-EMS, Can, Fra-EMS, Bel-(Ger+Neth) and Neth-Ger. These components have been ordered according to their "explained variance", that is, their sample variance divided by the sum of the sample variances of the individual exchange rate changes (the "total variance"). The explained variance is commonly used as a measure of importance of the principal components. It shows that the component dominated by the European currencies, the EMS component, is the most important one, explaining 77 percent of the total variance.

The remaining part in the estimation of the model concerns the estimation of the univariate GARCH models in (5.2) for each principal component. The results, as reported in table 5.3, are standard. Most importantly, they strongly reflect the presence of conditional heteroskedasticity. According to our model, this is the source of time-variation in the conditional variances as well as correlations of the individual exchange rate changes (see subsection 5.2.3).

5.4.3 Have Exchange Rates Become More Closely Tied?

Having estimated our multivariate GARCH model, we can now analyze how exchange rate correlations have evolved over the post-Bretton-Woods period. This is to answer the central question of the chapter, namely whether exchange rates have become more closely tied. Note that the conclusions will be in terms of nominal exchange rates. However, they are likely to hold for real exchange rates as well, because prices are fixed in the short run.

In figure 5.1 we plot the estimated correlations between several dollar exchange rates.[4] For the sake of exposition, we have smoothed the actual estimates by an equally weighted moving average using the estimates in the year before and the year after the week under consideration. Despite this smoothing, we still see that the correlations are not constant over time.

From figure 5.1 one may distinguish three remarkable periods for exchange rate correlations, roughly spanning the seventies, eighties and the nineties. The seventies are characterized by a decrease in correlation followed by an increase. The decrease may well be caused by the rather autonomous monetary and fiscal responses of governments to the 1974-1975 period of stagflation (see Krugman and Obstfeld (1991)). This policy imbalance, however, caused a steep depreciation of the U.S. dollar, so that Germany and Japan intervened heavily in the foreign exchange market in 1977-1978. Together with the inception of the EMS in 1979, this marks a period of greater coordination, causing the correlations to rise.

The eighties characterize a period of high correlations. This is confirmed by Bollerslev (1990), who finds that correlations between the European currencies were higher during the EMS period than before. In addition, we find that also intercontinental correlations were high. This is mainly caused by the huge swing in the dollar in the eighties. First, the dollar strongly appreciated partly due to the Volcker monetary contraction starting in 1979. In the second half of the eighties, coordinated actions such as the 1985 Plaza agreement brought the dollar down again. Moreover, the 1987 Louvre target zones may also explain the high correlations in the eighties.

The third remarkable period in figure 5.1 concerns the decrease in the correla-

[4] The estimates are based on the estimation results for the principal components in subsection 5.4.2, and the second relation in (5.3), which specifies the conditional variance of the exchange rate changes as a function of the conditional principal component variances.

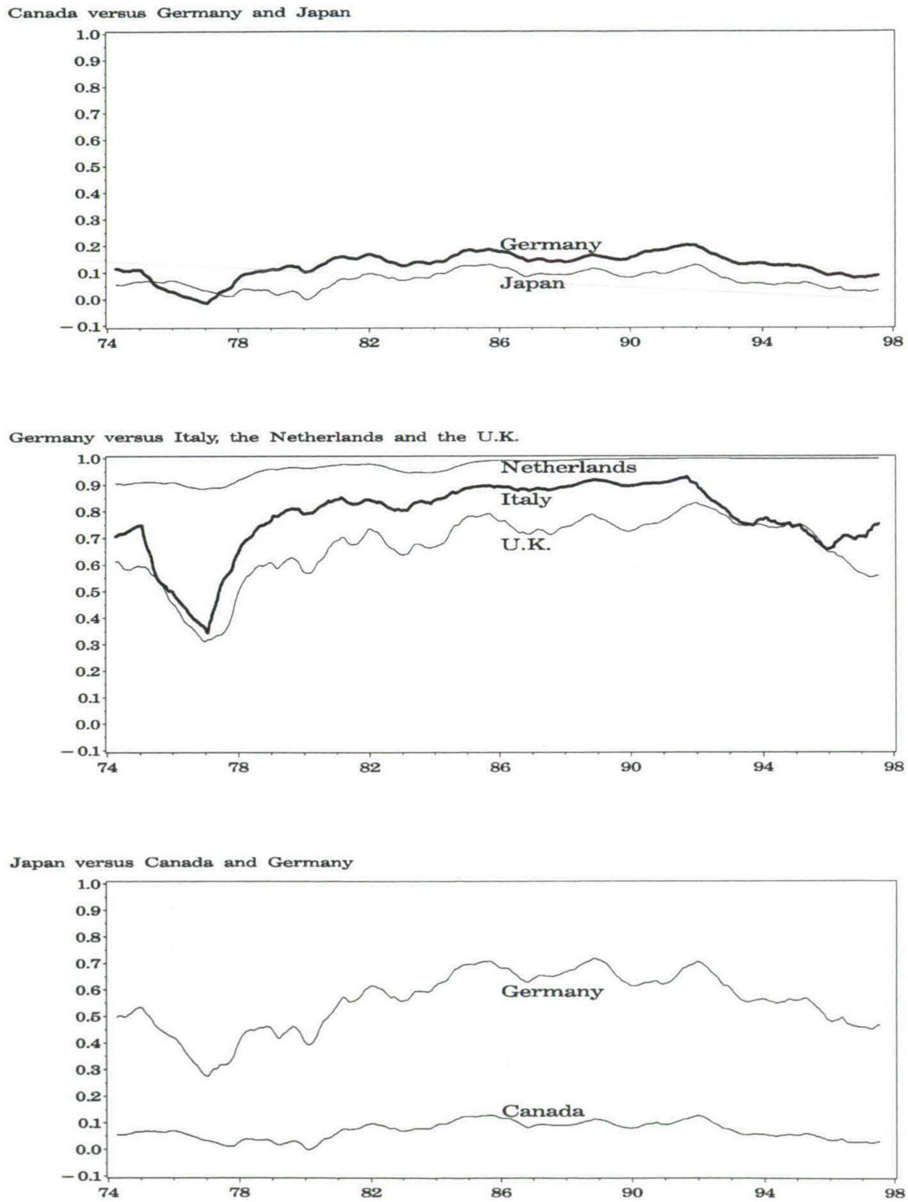


Figure 5.1: Smoothed estimated conditional correlations between dollar exchange rates

tions in the nineties. Hence, the main exchange rates have become more loosely instead of closely tied. At first sight, this may seem surprising, as it is often believed that the greater integration of financial markets has increased financial correlations. However, more integration also means that capital can move more freely, which can destabilize exchange rates. This happened in 1992 when the EMS collapsed, leading to a drop in several intra-EMS correlations, as shown by the middle graph of the figure. Furthermore, although European and American markets have become more integrated, Japan is still a relatively independent market. This may be the reason behind the fact that swings in the EMS - U.S. dollar rates have become shorter in the nineties, while the yen - dollar swings are still relatively long (see Chapter 4 for empirical evidence). These differences between the European currencies and the yen have also decreased the correlations in the nineties.

With the advent of European monetary unification (EMU), it is likely that the correlations between the participating European currencies will increase again at the end of the nineties. The upward tendency in the Germany - Italy conditional correlations after 1996 may be an indication of this. It will be interesting to analyze whether EMU also affects the correlations between the world's main currencies, namely, the U.S. dollar, yen and euro.

5.4.4 Diagnostics

The correlation analysis in the previous subsection was based on the multivariate GARCH model of subsection 5.2.1. In the remaining part of this section, we check empirically whether that model is appropriate for such an analysis. In the current subsection we examine whether it captures the features of the data described in subsection 5.4.1. In subsection 5.4.5 we compare the performance of our model with that of the Bollerslev (1990) and factor GARCH models.

To check the specification of our model, we analyze the normalized residuals. They are defined by $\widehat{\eta}_t = \widehat{V}_{t-1}\{\varepsilon_t\}^{-1/2} \cdot \widehat{\varepsilon}_t$, where $\widehat{V}_{t-1}\{\varepsilon_t\}^{-1/2}$ is the inverse of the lower triangular Cholesky decomposition of $\widehat{V}_{t-1}\{\varepsilon_t\}$ and ε_t is the exchange rate innovation $s_t - E_{t-1}\{s_t\}$. Table 5.4 presents several test results for them. The i-th column in this table concerns the i-th element of $\widehat{\eta}_t$. Unfortunately, we cannot attribute this element to one country, because $\widehat{\eta}_{it}$ is a linear combination of the country specific residuals $\widehat{\varepsilon}_{1t}, ..., \widehat{\varepsilon}_{it}$. We conclude from the first-order autocorre-

Table 5.4: Diagnostic statistics for normalized residuals

	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$	$i=7$	$i=8$
Autocorr. $\widehat{\eta}_{it}$: ρ_1	0.06* (0.03)	0.02 (0.03)	-0.04 (0.03)	0.05 (0.03)	0.00 (0.03)	-0.00 (0.03)	-0.02 (0.03)	-0.04 (0.03)
Autocorr. $\widehat{\eta}_{it}$: Q_{10}	19.83 [0.03]	9.16 [0.52]	10.69 [0.38]	5.42 [0.86]	8.29 [0.60]	18.19 [0.05]	13.33 [0.21]	17.24 [0.07]
Autocorr. $\widehat{\eta}^2_{it}$: ρ^s_1	0.01 (0.03)	0.02 (0.03)	-0.00 (0.03)	0.03 (0.03)	0.01 (0.03)	0.06* (0.03)	0.14* (0.03)	0.03 (0.03)
Autocorr. $\widehat{\eta}^2_{it}$: Q^s_{10}	11.65 [0.31]	4.05 [0.95]	1.37 [0.99]	2.88 [0.98]	6.52 [0.77]	13.04 [0.22]	26.81 [0.00]	3.71 [0.96]
Autocorr. $\widehat{\eta}_{it}\cdot\widehat{\eta}_{jt}$: $\overline{\rho^c_1}$	0.03 (0.03)	0.03 (0.03)	0.02 (0.03)	0.02 (0.03)	0.02 (0.03)	0.01 (0.03)	0.03 (0.03)	-0.01 (0.03)
Autocorr. $\widehat{\eta}_{it}\cdot\widehat{\eta}_{jt}$: $\overline{Q^c_{10}}$	17.00 [0.07]	10.67 [0.38]	6.09 [0.81]	7.77 [0.65]	9.84 [0.45]	11.06 [0.35]	7.41 [0.69]	14.91 [0.14]
Zero conditional correlation	10.77* [0.01]	2.76 [0.43]	3.07 [0.38]	4.26 [0.23]	7.05 [0.07]	3.64 [0.30]	5.65 [0.13]	4.94 [0.18]

Standard errors in parentheses and p-values in square brackets; * is significant at 5% level.
The vector of normalized residuals is $\widehat{\eta}_t = \widehat{V}_{t-1}\{\varepsilon_t\}^{-1/2}\cdot\widehat{\varepsilon}_t$, where $\widehat{V}_{t-1}\{\varepsilon_t\}^{-1/2}$ is the inverse of the lower triangular Cholesky decomposition of $\widehat{V}_{t-1}\{\varepsilon_t\}$ and ε_t is the exchange rate innovation $s_t - E_{t-1}\{s_t\}$. Hence, $\widehat{\eta}_{it}$ is a linear combination of $\widehat{\varepsilon}_{1t}, ..., \widehat{\varepsilon}_{it}$, so that, in contrast to $\widehat{\varepsilon}_{it}$, $\widehat{\eta}_{it}$ does not directly correspond to one country.
All autocorrelation statistics have been defined below table 5.1, although the standard error of ρ_1 and the value of Q_{10} are no longer corrected for heteroskedasticity.
The test for zero conditional correlation between $\widehat{\eta}_{it}$ and the other seven $\widehat{\eta}_{jt}$ is similar to the constant correlation test of table 5.1. However, now we also test for a zero intercept in the regressions involved. Hence, the critical value is based on a χ^2_3 instead of χ^2_2 distribution.

lations and the Box-Pierce statistics Q_{10} that there is no evidence of remaining autocorrelation in the normalized residuals.

Secondly, the measures for remaining autocorrelation in the squared changes and the cross products show no reason to extend the variance specification of the model.

The final test in table 5.4 also concerns the variance specification, as it checks whether the normalized residuals are conditionally uncorrelated. This is done by regressing the cross products of the normalized residuals at time t on the vector $(1, t, t^2)'$ and testing whether all three regression coefficients are zero. The difference with the test for short-run autocorrelation in the cross products, as discussed in the previous paragraph, is that the current test has more power against long-run autocorrelation. Moreover, it also tests whether the cross products have mean zero. The results in table 5.4 again show no serious evidence of misspecification.

It is interesting to observe that the test for zero conditional correlations of the normalized residuals is similar to the test for constant conditional correlations of the exchange rate changes in subsection 5.4.1. The latter test was clearly rejected, but the test on the residuals of our model is not. Apparently, our model is able to describe the time-varying pattern in the conditional exchange rate correlations quite well. This is the main reason why we prefer our model over the Bollerslev (1990) constant conditional correlations model, as our study is focused on the time-variation in exchange rate correlations. In this respect, our model is also preferable over the usual factor GARCH model, which would be 1-factor GARCH for our data, as argued below. Although that model captures some time-variation in conditional correlations, it does not explain it completely, as six out of eight zero-conditional-correlation statistics are significant.[5]

5.4.5 Goodness of Fit

In the introduction we have claimed that our multivariate GARCH model provides a good fit for the conditional exchange rate variances and correlations, at least compared to the Bollerslev (1990) and the usual factor GARCH models with much less than $I=8$ factors. In this subsection we provide evidence for that. We also examine the reasons behind the outperformance by analyzing the variance and correlation fits separately.

To measure the goodness of fit of the models, we use the multivariate normal log-likelihood with conditional mean and variance as estimated by the different models. The "total fit" column of table 5.5 contains the results. It shows that the log-likelihood of our model, -8,817, is better than the one of the constant conditional correlations model of Bollerslev (1990), which is -10,624.

To compare our model with the usual factor GARCH model, we first have to choose the usual number of factors, or principal components, K. The commonly used Kaiser-Guttman rule states that one should select only principal components that have a larger variance than the average variance of the exchange rate changes (see Bartholomew (1987)). For our data, this rule leads to $K=1$, as only the variance of the EMS component (11.60, see table 5.2) exceeds the average variance

[5]The zero-conditional-correlation tests for 1-factor GARCH are 21.34 [with p-value 0.00], 10.36 [0.02], 10.77 [0.01], 19.85 [0.00], 26.26 [0.00], 5.83 [0.12], 4.84 [0.18] and 21.52 [0.00]; see the note below table 5.4 for the definition of these test.

Table 5.5: Quality of various multivariate GARCH models

Model	TOTAL FIT			VARIANCE FIT		CORREL. FIT	
	log-lik.	change	LR	log-lik.	change	log-lik.	change
Univar. GARCH	−15,864	0	–	−15,864	0	0	0
Const. cond. corr.	−10,624	5420	10,840*	−15,864	0	5240	5240
1-factor GARCH	−10,315	309	–	−16,078	−214	5763	523
2-factor GARCH	−10,248	67	134*	−16,037	41	5789	67
3-factor GARCH	−10,145	103	206*	−16,013	24	5867	103
4-factor GARCH	−9,836	309	618*	−15,958	55	6122	309
5-factor GARCH	−9,796	40	80*	−15,916	42	6120	40
6-factor GARCH	−9,579	217	434*	−15,937	−19	6358	217
7-factor GARCH	−9,237	342	684*	−15,936	1	6699	342
Our model	−8,817	420	840*	−15,935	1	7118	420

A * denotes significance at the 5% level.
The quality measure we use is the log-likelihood based on a normally distributed vector of exchange rate changes with conditional mean and variance as estimated by the different models. In the "total fit" column, the full estimated conditional variance matrix is used to compute this log-likelihood. For the "variance fit" column, the conditional correlations have been substituted by zero. The "correlation fit" column is the difference between the "total fit" and "variance fit" columns.
The "total fit" column also contains the likelihood ratio (LR) for the model against the previous one, if the model includes the previous one as a special case.
"Univar. GARCH" is the model that imposes diagonality of the conditional variance matrix, so that the moments can be estimated by eight univariate GARCH procedures.
"Const. cond. corr." denotes the Bollerslev (1990) model with constant conditional correlations. It is estimated in two steps. First, we estimate eight univariate GARCH models, and then we derive the conditional correlation estimates.
For the K-factor GARCH models, the conditional mean and variance follow from the multivariate second estimation step (see section 5.3.2 and footnote 2). For parsimony, we assume that the covariance matrix of the exchange rate specific changes v_t in (5.6), $V\{v_t\}$, is diagonal, as in Diebold and Nerlove (1989).

of 1.89. This is in line with the choice of Diebold and Nerlove (1989), who use about the same exchange rates.

Table 5.5 demonstrates that our model is preferable over the 1-factor GARCH model, which has a log-likelihood of -10,315, as the likelihood ratio is 2,996.[6] Hence, we conclude that our model indeed provides a better fit than the popular constant conditional correlations and 1-factor GARCH models. Note that our model also significantly outperforms the other factor GARCH models, as the likelihood ratios in table 5.5 show.

[6]The K-factor GARCH model is nested in our model, as it follows after restricting the last $K-I$ columns of the matrix of factor loadings in (5.6), Λ, to zero.

In the remaining part of this subsection, we investigate the reasons for this outperformance. We first analyze the variance fit and then the correlation fit.

To measure the variance fit, we remove the correlation effects from the log-likelihood by substituting the off-diagonal elements in the estimated conditional variance matrices by zero. The "variance fit" column of table 5.5 gives these zero-correlation log-likelihoods. Our model (-15,935) somewhat underperforms the constant conditional correlations model (-15,864). This is not surprising. The variance fit of the latter model is entirely based on univariate GARCH estimations for each exchange rate change and the univariate estimations only have to fit the conditional variance process, while our model is mainly designed to give a good description of the correlation process.

Table 5.5 also shows that our model outperforms the usual 1-factor GARCH model in terms of variance fit. The reason is that the first principal component is only a single combination of exchange rate changes, and one cannot expect that this would lead to good variance estimates for all exchange rate changes individually.[7] A good variance fit requires at least five principal components, as table 5.5 shows. The relevance of the fifth component, the one dominated by Canada, is shown by figure 5.2. For $K = 4$, the conditional variance estimates for the Canadian dollar are almost constant, while only inclusion of the Can component leads to a time-variation pattern that one also finds for univariate AR(1)-GARCH(1,1) on the Canadian dollar.

The correlation fit is the second reason for the relatively good fit of our model. It is measured by the difference between the full and the zero-correlation log-likelihoods, and it is reported in the "correlation fit" column of table 5.5. It is clear that our model outperforms the constant conditional correlations model. This again supports the conclusion that the assumption of constant conditional correlations is too restrictive for our data.

[7]It is interesting to observe that the lack of variance fit of the 1-factor GARCH model is hidden by the full log-likelihood, that is, the quality measure including the conditional correlations, which we have used at the beginning of this subsection. Recall that the full log-likelihood is -10,315, which is much greater than the sum of the log-likelihoods obtained from eight independent univariate AR(1)-GARCH(1,1) models for the exchange rate changes, which is -15,864. Hence, one is tempted to conclude that the 1-factor GARCH model is to be preferred; this is also what Diebold and Nerlove (1989) claim. However, the huge increase in the log-likelihood is entirely due to a better fit of the conditional correlations, and the log-likelihood is very sensitive to that (see also footnote 8). Hence, the log-likelihood of the factor model including the correlations can be a misleading indicator for the quality of the variance fit.

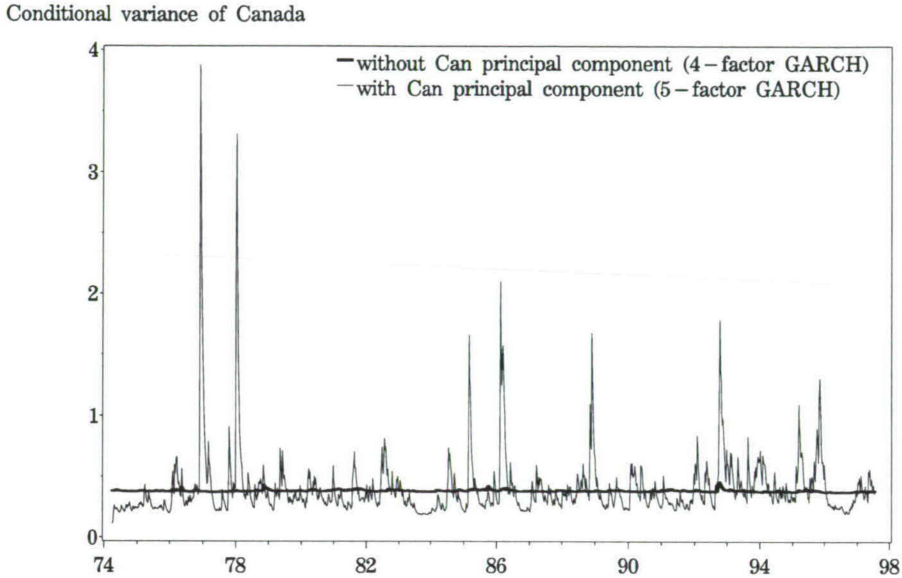


Figure 5.2: Effect of Can principal component on the estimated conditional variance of Canada

Table 5.5 also demonstrates that our model provides a better correlation fit than the 1-factor GARCH model. Moreover, it also outperforms the factor GARCH model with five factors, the number of factors that is at least needed for an acceptable variance fit. Although the final three components do not improve the variance fit, they do yield a better correlation fit. In fact, adding the last component increases the log-likelihood by 420, which is highly significant.[8] This can be attributed to a better fit of the time-variation in the conditional correlation between the Netherlands and Germany, as figure 5.3 demonstrates. Only the inclusion of the last component allows the factor GARCH model to capture that since the mid eighties the monetary policy of the Dutch central bank is mainly

[8]This huge significance (likelihood ratio is 840) is due to the great sensitivity of the log-likelihood to the correlation fit. This is also the reason why $K=1$ at first sight seems to be much better than eight univariate AR(1)-GARCH(1,1) estimations on the individual exchange rates, as shown in footnote 7. It also explains why Bollerslev (1990) gets a highly significant likelihood ratio test of almost 2,000 when testing for zero correlations in his multivariate GARCH model.

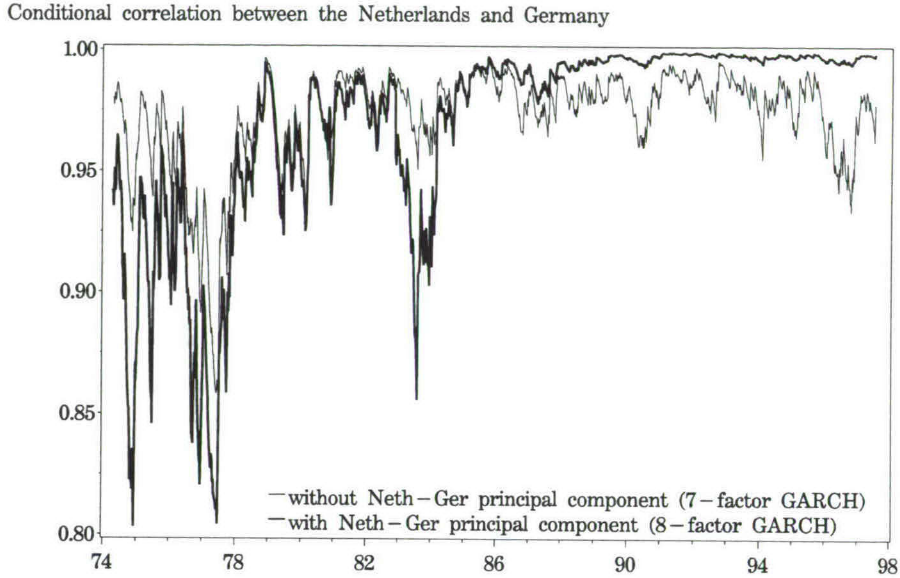


Figure 5.3: Effect of Neth-Ger principal component on the estimated conditional correlation between the Netherlands and Germany

attributed to keeping the guilder-mark rate stable, so that both currencies move more closely together than before.

In summary, the conclusion from this subsection is that our model results in a better fit than two popular multivariate GARCH models, namely the Bollerslev (1990) model and the usual factor GARCH model. This holds especially for the correlations, which we are particularly interested in.

5.5 Conclusion

In this chapter we analyze exchange rate correlations over time. For that, we introduce a new multivariate GARCH model. It describes the exchange rate changes indirectly through their principal components and assumes that the conditional variances of the components govern the conditional exchange rate correlations. We show that this is quite realistic, both from an economic and empirical point of view. Moreover, the indirect approach implies that the model is very easy to estimate, as it only requires several univariate GARCH estimations to estimate the full multivariate model.

The empirical results show that the model provides a better fit than existing models. First, it outperforms the popular constant conditional correlations model of Bollerslev (1990) with respect to the correlation fit. This is not surprising, as the data show clear evidence of time-variation in the conditional correlations and only our model can capture that. Second, our model provides a better variance and correlation fit than usual factor GARCH models. This is explained by the fact that our model can be viewed as a factor GARCH model with the maximum number of factors and that the factors neglected in usual factor GARCH contain important information for exchange rate variances and correlations.

Given the outperformance qua fit, we use our model to analyze the correlations between eight U.S. dollar exchange rates over the post-Bretton-Woods period. We find that these correlations were highest in the eighties and then decreased in the nineties. Hence, exchange rates have become more loosely instead of closely tied. This originates from the EMS crash in 1992, making several European exchange rates less correlated. Moreover, the EMS - yen correlations have decreased because of the combination of more stable EMS - U.S. dollar rates and a long swing in the yen - dollar rate.

So far, we have concentrated on GARCH in a multivariate setting. However, it is important to realize that our indirect approach via the principal components is not restricted to GARCH. In fact, any univariate model for the principal components can be used to derive a practical multivariate model. This offers a wide range of applications of our approach. For instance, when analyzing stock or bond return correlations, one can take account of asymmetric volatilities, GARCH-in-mean effects and other deviations from standard GARCH (see Bollerslev et al. (1992)). Furthermore, our approach can form the basis for multivariate extensions of other volatility models, such as stochastic volatility, regime-switching GARCH and fractionally integrated GARCH. This is left for future research.

Appendix

5.A Our Model is a Special Factor GARCH Model

In this appendix we demonstrate that our model of subsection 5.2.1 is a factor GARCH model with as many factors, K, as exchange rates, I. For that, we first define what we actually mean by the K-factor GARCH model. As in the main text, we concentrate on the conditional mean and variance of exchange rate changes. The final factor GARCH specification of these moments is derived in two stages.

To obtain the first factor GARCH formulation, we split the vector of exchange rate changes s_t into

$$s_t = \mu_t + \varepsilon_t, \tag{5.8}$$

where $\mu_t = E_{t-1}\{s_t\}$ and ε_t is the innovation. The central idea of the factor model is that ε_t has a systematic and an unsystematic part, where the systematic part is a linear combination of K unobserved factors φ_{kt}:

$$\varepsilon_t = \Lambda\varphi_t + \upsilon_t, \tag{5.9}$$

where $\varphi_t = (\varphi_{1t}, \ldots, \varphi_{Kt})'$ is the K-vector of common factors, Λ is the $I \times K$ full-column-rank matrix of factor loadings, and υ_t denotes the unsystematic, exchange rate specific change. We assume that $E_{t-1}\{\varphi_t\} = 0$ and $E_{t-1}\{\upsilon_t\} = 0$ to ensure $E_{t-1}\{\varepsilon_t\} = 0$. Moreover, let $V_{t-1}\{\varphi_t\}$ denote the time-varying conditional variance of φ_t.[9] Let $V_{t-1}\{\upsilon_t\}$ be the variance of υ_t, which we assume constant over time ($V_{t-1}\{\upsilon_t\} = V\{\upsilon_t\}$), as in Engel et al. (1990). Finally, we have $Cov_{t-1}\{\varphi_t, \upsilon_t\} = 0$.

The main effect of the factor model is that it puts structure onto the innovation ε_t. However, as in Engle et al. (1990), the factor idea can also be used to specify the expected exchange rate changes μ_t. This makes μ_t the sum of a systematic part, which is attributed to the factors, and an unsystematic part. More formally,

$$\mu_t = \Lambda\mu_t^{\varphi} + \mu^{\upsilon}, \tag{5.10}$$

[9] Note that we do not impose diagonality of $V_{t-1}\{\varphi_t\}$. Diagonality has been commonly used in the literature to help identify Λ. Later on, we will introduce another, very convenient way to identify Λ.

where the systematic part is a linear combination of a K-vector of common sources of expected depreciation, μ_t^φ, and the unsystematic part is an I-vector of exchange rate specific expected depreciations, which we assume constant over time ($\mu_t^\upsilon = \mu^\upsilon$).

Specifications (5.8), (5.9) and (5.10) lead to the first formulation of the factor GARCH model in terms of the moments of interest:

$$\begin{aligned} E_{t-1}\{s_t\} &= \Lambda\mu_t^\varphi + \mu^\upsilon \\ V_{t-1}\{s_t\} &= \Lambda V_{t-1}\{\varphi_t\}\Lambda' + V\{\upsilon_t\}. \end{aligned} \tag{5.11}$$

This holds for all $K \in \{1, \ldots, I\}$. Note that for $K = I$, the case we are particularly interested in, the parameters μ^υ and $V\{\upsilon_t\}$ are zero, because in that case ε_t (μ_t) is one-to-one related to φ_t (μ_t^φ).

The factor model in its current format cannot be estimated because of two problems. The first one is that the systematic and unsystematic innovations, φ_t and υ_t, are not observed separately, so that the parameters are, in general, not directly estimable. The second problem is caused by a rotational indeterminacy in the definition of the factors, which makes Λ unidentified. We now solve both problems in turn, so as to derive the second factor GARCH moment specification.

As shown by Engle et al. (1990), the first problem can be solved by substituting the unobserved factors φ_t by an expression based on an observed K-vector that is closely related (but not equal) to the factors in a sense that is explained at the end of footnote 11. Similar to many other papers (for instance, see Ng et al. (1992)), we take K principal components of s_t to form this factor representing vector, and we assume that they are conditionally uncorrelated and that each of them follows an AR(1)-GARCH(1,1) model. Hence, the factor representing vector is a K-dimensional subvector of f_t, the vector of all I principal components described by (5.1) and (5.2). For simplicity of notation, let us denote this subvector of f_t also by f_t, and let W also denote the $I{\times}K$ full-column-rank matrix of component weights that defines the subvector by $f_t = W's_t$. Using (5.11), this implies that

$$\begin{aligned} E_{t-1}\{f_t\} &= W'\Lambda\mu_t^\varphi + W'\mu^\upsilon \\ V_{t-1}\{f_t\} &= W'\Lambda V_{t-1}\{\varphi_t\}\Lambda' W + W'V\{\upsilon_t\}W. \end{aligned} \tag{5.12}$$

Since $W'\Lambda$ is invertible, we can solve μ_t^φ and $V_{t-1}\{\varphi_t\}$ from these equations and substitute the results in (5.11). This gives

$$E_{t-1}\{s_t\}=\Lambda(W'\Lambda)^{-1}E_{t-1}\{f_t\}-\Lambda(W'\Lambda)^{-1}W'\mu^{\upsilon}+\mu^{\upsilon} \tag{5.13}$$
$$V_{t-1}\{s_t\}=\Lambda(W'\Lambda)^{-1}V_{t-1}\{f_t\}(\Lambda'W)^{-1}\Lambda'-\Lambda(W'\Lambda)^{-1}W'V\{\upsilon_t\}W(\Lambda'W)^{-1}\Lambda'+V\{\upsilon_t\}.$$

The main difference with (5.11) is that (5.13) contains only parameters related to the unsystematic innovation υ_t, not related to the factors φ_t, as the observable f_t has taken the role of φ_t. Therefore, using the principal components has solved the first problem.

The second problem with (5.11) is caused by a rotational indeterminacy in the unobserved factors, so that Λ is not identified. That is, if a certain combination of Λ, μ_t^{φ} and φ_t gives the true conditional moments of s_t, then, for any invertible $K{\times}K$-matrix Q, the oblique rotations ΛQ, $Q^{-1}\mu_t^{\varphi}$ and the oblique factors $Q^{-1}\varphi_t$ yield the same conditional moments.

Formula (5.13) shows this problem again. Since Λ only occurs in the combination $\Lambda(W'\Lambda)^{-1}$, it is only identified if we can derive its $I{\cdot}K$ unknown elements from a particular value of $\Lambda(W'\Lambda)^{-1}$, say A. However, this is impossible, since there are only $I{\cdot}K{-}K^2$ independent equations in $\Lambda(W'\Lambda)^{-1}=A$.[10] Therefore, we need K^2 normalizing restrictions on Λ. Considering (5.13), it is very convenient to use $W'\Lambda=I_K$, where I_K is the $K{\times}K$ identity matrix.[11] We will see below that this normalization is crucial for proving that our model is an I-factor GARCH model.

Having solved both problems, we can present the second and final factor

[10] The system $\Lambda(W'\Lambda)^{-1}=A$ is equivalent to $(I_I-AW')\Lambda=0$, where I_I is the identity matrix of dimension I. To compute the rank of I_I-AW', we first note that AW' is idempotent, since $W'A=I_K$. Hence, the rank of I_I-AW' is $r(I_I-AW')=I-r(AW')$. Moreover, $r(AW')=K$, since both A and W' have rank K. Therefore, the rank of I_I-AW' is $I-K$, so that the system $(I_I-AW')\Lambda=0$ contains exactly $(I-K)\cdot K$ independent equations.

[11] This normalization has three interesting characteristics. First, it directly reduces the number of free parameters, which makes estimation simpler. For instance, for $K=7$ and $I=8$, it implies that only seven factor loadings have to be estimated instead of 56.
The second characteristic of our normalization is that it is necessary and sufficient. This is in contrast with the sufficient identifying restrictions employed by Sentana (1992) and King, Sentana and Wadhwani (1994), who impose $V\{\varphi_t\}=I_K$ for the conditional variance of the factors to identify Λ (except for column sign).
Finally, our normalization explains in what sense the principal components are "closely related" to the factors. Using $W'\Lambda=I_K$ in the conditional variance of f_t, which is $V_{t-1}\{f_t\}=W'\Lambda V_{t-1}\{\varphi_t\}\Lambda'W+W'V\{\upsilon_t\}W$, shows that the conditional variance of each component f_{kt} is perfectly correlated with that of the k-th factor φ_{kt}. This is why the f_{kt} are called "factor representing portfolios" in Engle et al. (1990).

GARCH formulation, which is commonly used in the literature:

$$
\begin{aligned}
E_{t-1}\{s_t\} &= \gamma + \Lambda E_{t-1}\{f_t\} \\
V_{t-1}\{s_t\} &= \Omega + \Lambda V_{t-1}\{f_t\}\Lambda',
\end{aligned} \tag{5.14}
$$

where $\gamma = (I_I - \Lambda W')\mu^v$ and $\Omega = V\{v_t\} - \Lambda W' V\{v_t\} W \Lambda'$ are the time-constant parts in the mean and variance, respectively. Note that these parts are zero in case of $K = I$, because then μ^v and $V\{v_t\}$ are zero.

Although some similarities with our model of section 5.2 have already become clear, it may not yet be clear that our model exactly equals the factor GARCH model for $K = I$. The final link is provided by our factor GARCH normalization $W'\Lambda = I_K$. In case of $K = I$, this normalization and the orthogonality of W imply that $\Lambda = (W')^{-1} = W$. Therefore, relation (5.14), where γ and Ω are zero because of $K = I$, is the same as the second part of our model, that is, formula (5.3). Because the model for the I principal components is also the same, our model is indeed a special factor GARCH model in which the number of factors equals the number of exchange rates.

Chapter 6

Why is it so Difficult to Find an Effect of Exchange Rate Risk on Trade?

It is commonly argued that exchange rate risk depresses international trade. However, the large literature on this subject has not yet provided conclusive evidence. This chapter analyzes why it is so difficult to obtain a clear answer from time series analyses. We use data on bilateral aggregate U.S. exports to the other G7 countries. The results show that export decisions are mostly affected by the exchange rate about one year later. The riskiness of the exchange rate at such a long horizon appears fairly constant over time with only short-term fluctuations. This makes it difficult to discover the true effect of exchange risk on trade from the limited time series data that are typically available.

6.1 Introduction

It is commonly claimed that exchange rate risk has a negative effect on international trade. The standard argument is that greater exchange risk increases the riskiness of trade profits, leading risk averse traders to reduce trade.

Because of this widespread view, the effect of exchange rate risk on trade has been important for various economic policy discussions. For instance, it is important for the choice between a fixed and floating exchange rate regime. In this respect, it was used as one of the main economic arguments for European Monetary Unification (see EU Commission (1990)).

Also within a floating regime the effect of exchange risk on trade is important. For example, it provides a rationale for foreign exchange interventions, such as those following the 1987 Louvre Accord. After all, one of the motives for intervention is to reduce exchange rate fluctuations, because exchange rate risk is assumed to have an adverse effect on trade (see Edison (1993) and Almekinders and Eijffinger (1991)). This theoretical argument is supported by Bayoumi and Eichengreen (1998), who demonstrate empirically that trade links between countries encourage foreign exchange intervention.

Given its economic relevance, the effect of exchange rate risk on international trade has attracted many researchers in international economics. The voluminous theoretical and empirical literature, however, has not yet provided conclusive evidence, despite the widespread view of a negative effect.[1] In this chapter we try to explain why it is so difficult to find a clear effect.

We first empirically re-examine the effect of risk on trade for our data set, which concerns monthly bilateral aggregate U.S. exports to the other G7 countries from 1978 to 1996. The chapter pays special attention to several methodological issues. For instance, compared to existing studies, we reduce measurement error in the crucial exchange risk measure by using daily exchange rates to quantify multi-month-ahead real exchange risk. Moreover, to enhance the dynamic structure of our distributed lag model and to determine which exchange risk horizon is relevant for goods traders, we introduce a new parsimonious lag structure using the Poisson probability (mass) function to distribute the total effect of a regressor over time. Both methodological issues will be discussed in more detail later on in this introduction. Our results on the effect of real exchange risk on exports confirm the ambiguity found in the literature.

Next, we address the main focus of the chapter, that is, we analyze why it is so difficult to find a clear effect. We concentrate on time series analyses, as they are used in the vast majority of existing empirical studies. The estimates show that export decisions are mostly affected by the exchange rate distribution about one year later. The riskiness of the exchange rate at such a long horizon appears fairly constant over time with only short-term fluctuations. This makes it so difficult to discover the true effect of risk on trade from the limited time series data that are typically available.

[1]See the survey article by Côté (1994) and the references therein. More recent theoretical papers are Broll and Eckwert (1997) and Bacchetta and Van Wincoop (1998). Recent empirical work includes Caporale and Doroodian (1994), Qian and Varangis (1994), Arize (1995) and Fountas and Aristotelous (1999).

The remaining part of this introduction explains the contribution of this chapter to the literature in more detail.

In general, there can be several reasons for the ambiguity found in the empirical literature on the effect of exchange rate risk on trade. Here, we discuss three of them (see Côté (1994) for additional reasons). First, the effect may indeed be absent, for instance, because firms can avoid all exchange risk by hedging. However, Wei (1999) finds no support for the hedging argument. The absence of any effect would also be in contrast with the widespread view of a negative effect.

A second reason, stressed by Bini-Smaghi (1991), is that the empirical tests may be subject to methodological problems. One issue concerns the measurement of exchange rate risk, which is assumed to be equal to exchange rate volatility or variability, as usual in the trade literature. Quite surprisingly, the measurement of risk has received only moderate attention in the trade literature, despite the central role of this variable. Many authors use the moving standard deviation of the past, say, 24 monthly exchange rate changes for simplicity.[2] Others use a generalized autoregressive conditional heteroskedasticity (GARCH) model, given the popularity of this model to capture the strong volatility clustering in high-frequency time series.[3] We demonstrate that both measures have conflicting implications for the evolvement of risk over time, at least concerning the long-term risk that is relevant for goods traders. The moving standard deviation measure implies that exchange rate shocks persist in risk for a considerable period of time (24 months in our example), suggesting high serial correlation in risk. The GARCH measure, on the other hand, yields a low or even zero persistence of shocks in monthly risk, suggesting low or no serial correlation in risk. To solve this contradiction we use an alternative risk measure based on Merton (1980) and Andersen and Bollerslev (1998). Instead of taking monthly squared changes, we compute monthly exchange rate volatilities by cumulating squared daily changes in the month. Then we estimate an autoregressive model of order two on the monthly (estimated) volatilities, and we use the AR(2) forecasts to define multi-month-ahead exchange rate risk. We show that this measure describes the serial

[2]For instance, Cushman (1983, 1986), Klein (1990), Chowdury (1993), Arize (1995) and Fountas and Aristotelous (1999).

[3]See Bollerslev, Chou and Kroner (1992) for an overview of GARCH. GARCH risk measures are used in Pozo (1992), Kroner and Lastrapes (1993), Caporale and Doroodian (1994) and Qian and Varangis (1994), among others.

correlation in risk better than the two measures commonly used in the trade literature. Hence, our measure yields a reduction in measurement error for the crucial exchange risk variable, making the estimated effect of risk on exports more accurate.

Another methodological issue we address concerns the dynamic specification of the trade model. We employ a distributed lag model and introduce a new way to impose structure on the lag coefficients. Our method separates the total effect of a regressor from the distribution of the effect over time. The latter part appears to be a probability function, which can be freely chosen. For convenience, we use the Poisson probability function. This lag structure turns out to be more appropriate than the commonly used geometric and polynomial lags, because the Poisson lag structure can capture hump shaped lag patterns and it avoids sign-switching of the estimated lag coefficients. The estimates for the Poisson parameters show that foreign income has the largest effect on domestic exports after about one quarter, while for the exchange rate this occurs only after about one year. Such time lags underscore the importance of allowing for dynamics in trade equations.

In summary, we take account of some important methodological issues that may explain the ambiguous results in existing trade studies. Nevertheless, we still find no clear effect of real exchange rate risk on trade. Hence, methodological problems are no sufficient explanation.

A third reason for the empirical ambiguity may come from the characteristics of exchange risk. Gagnon (1993) shows in a simulation experiment that the exchange risk level currently observed among industrial countries is too low to yield statistically detectable effects on trade. Our work is complementary to Gagnon (1993) in the sense that we study the time-variation instead of the level of risk. We empirically demonstrate that the time-variation in risk at the long horizon relevant for goods traders is rather low and that deviations from average risk do not persist long. Therefore, even if risk affects exports, the effect captures only a minor part of the time-variation and the long-term swings in exports; other shocks to exports are likely to overshadow any risk effect. We conclude that the two characteristics of long-term real exchange rate risk just mentioned make it difficult to discover the true effect of risk on exports from the limited time series data that are typically available.

The chapter is organized as follows. In section 6.2 we use an economic model to introduce the variables we think are important for the empirical work. Section 6.3 describes how we measure these variables. Given the importance of the exchange risk variable, we explain its measurement in detail in subsection 6.3.2. Section 6.4 presents the empirical model with special attention to the Poisson lag structure in subsection 6.4.2. In section 6.5 we report the empirical results and explain why we think it is so difficult to find the true effect of risk on trade from time series analyses. Section 6.6 concludes.

6.2 Economic Model

In this section we develop an economic model for the determination of exports. It provides a motivation for the choice of explanatory variables in the econometric model for U.S. exports that will be used later on.

The economic model is based on the popular two-country imperfect substitutes model (see Goldstein and Khan (1985)), which considers domestic exports and goods produced abroad as imperfect substitutes. The extension we make to the standard imperfect substitutes model is that we explicitly account for the lag between the time of the trade decision and the time of the actual trade flow and payment. This time lag is an important characteristic of international trade, as Goldstein and Khan (1985) and Sawyer and Sprinkle (1997) argue. Its existence implies that exchange rate risk can affect trade, as the exchange rate needed to convert foreign currency payments is unknown at the time of decision making.

Let t denote the time (month) of observing a nominal export flow X_t from the home to the foreign country, expressed in domestic currency. Exports are, supposedly, the result of a contract signed l months earlier, stating both the export quantity Q_{xt} and price P_{xt}. For simplicity, we assume that the price is specified in the home currency, so that $X_t = Q_{xt}P_{xt}$.[4]

Our focus variable is (the logarithm of) the real value of exports, using the

[4]The model can be extended to allow for invoicing in foreign currency as well. In that case, X_t also depends on the contemporaneous nominal exchange rate, which converts the foreign currency invoiced part of exports into domestic currency. It can be shown that the collection of export determinants in the final model equation (6.5) should then be extended by the contemporaneous real exchange rate. We can avoid this extra complexity, because in the empirical part of the chapter we study U.S. exports and these are almost completely invoiced in U.S. dollars (see Page (1981) for empirical evidence).

price P_t of domestically produced goods as deflator:

$$x_t = q_{xt} + p_{xt}, \tag{6.1}$$

where $x_t = \log(X_t/P_t)$, $q_{xt} = \log(Q_{xt})$ and $p_{xt} = \log(P_{xt}/P_t)$. We concentrate on the value x_t rather than the quantity q_{xt}, as is often done in the literature, because we study bilateral exports for which x_t is directly observable, while there are no observations on the bilateral prices needed to derive q_{xt} from x_t.

The determinants of x_t follow from the assumptions regarding export supply and demand. Supply is an unknown function q_x^s of only the price of exports relative to the price of domestic output in month t:[5]

$$q_{xt}^s = q_x^s(p_{xt}). \tag{6.2}$$

Foreign demand for domestic exports depends on two components. First, we suppose that it depends on real foreign income. Since the trade decision is made l months before the actual trade flow in month t, we use (the logarithm of) lagged real foreign income y_{t-l}^*.

The second determinant of foreign demand is the price of traded goods relative to the price P_t^* of foreign produced goods, both in foreign currency. Since the traded goods are invoiced in domestic currency, this relative price can be expressed as $P_{xt}/S_t \cdot 1/P_t^*$, where S_t is the nominal (spot) exchange rate, that is, the domestic currency price of one unit of foreign currency. In logarithms, the relative price equals $p_{xt} - s_t$, where $s_t = \log(S_t P_t^*/P_t)$ is the real exchange rate.[6] Although it is implicitly assumed that P_t and hence p_{xt} are perfectly forecastable at time $t-l$, such an assumption is not realistic for s_t, at least not for floating exchange rates. Hence, we account for the randomness of s_t at the time $t-l$ the trade decision is made. As usual in the trade literature, we assume that the mean and standard deviation of s_t, conditional on information I_{t-l} available at time t–l, are sufficient to capture the effects of exchange rates on export demand.[7]

[5] We take the price level P_t of the month of the export flow, month t, to deflate the export price, because we assume that the exporter receives payment in the same month as the delivery of the goods. This assumption is quite reasonable, as Stokman (1995) reports that payments peak in the month of delivery.

[6] Throughout this chapter the definitions of S_t and s_t are different from the ones used in previous chapters; there, S_t was in logarithms and s_t was the percentage nominal exchange rate depreciation.

[7] For simplicity, we abstract from the existence of a forward market to hedge exchange rate

Combining the income and price components just discussed, we specify the demand for domestic exports as

$$q_{xt}^{d} = q_{x}^{d}\left(y_{t-l}^{*}, E_{t-l}\{p_{xt}-s_t\}, V_{t-l}^{1/2}\{p_{xt}-s_t\}\right), \tag{6.3}$$

where E_{t-l} and $V_{t-l}^{1/2}$ denote the mean and standard deviation conditional on I_{t-l}.

The market for domestic exports is in equilibrium if

$$q_{xt} = q_{xt}^{s} = q_{xt}^{d}. \tag{6.4}$$

Solving (6.2)-(6.4) for p_{xt} and q_{xt} and substitution in (6.1) then yields

$$x_t = x\left(y_{t-l}^{*}, E_{t-l}\{s_t\}, V_{t-l}^{1/2}\{s_t\}\right). \tag{6.5}$$

Hence, the determinants of real (domestic output) exports are real foreign income (with an expected positive effect), the expected real exchange rate level (positive effect) and real exchange rate risk (unknown effect). The inclusion of income and the real exchange rate level is standard in trade models, in particular models that are also based on the imperfect substitutes model (see Goldstein and Khan (1985)). The extra real exchange risk term in (6.5) originates from the lag between the contract time $t-l$ and the time t of delivery and payment and from the fact that foreign demand depends on the exchange rate, which is unknown at time $t-l$.

6.3 Data Characteristics

In this section we first describe the data we use to measure the variables in (6.5), as these are the variables that will appear in the econometric model later on. We then pay specific attention to the measurement of the conditional standard deviation $V_{t-l}^{1/2}\{s_t\}$. Finally, we study the stationarity of the variables.

6.3.1 Data

The data is monthly bilateral aggregate U.S. exports to the six other G7 countries, namely Canada, France, Germany, Italy, Japan and the U.K.. We use bilateral

risk. Because the forward exchange rate is highly dependent on the mean and standard deviation of the future spot rate (see Viaene and De Vries (1992)), which we both take account of, the benefits from including the forward rate are likely small.

instead of the often used multilateral data to avoid the difficult construction of multi-country explanatory variables. Moreover, by considering several export flows that are selected in a rather natural manner we can provide some insight into the robustness of our results. The fact that we use aggregate instead of product-specific trade data is not important for the focus of the chapter, as shown in subsection 6.5.2.

The export time series span January 1978 through November 1996, leading to 227 monthly observations. For the other variables we require longer series because of the lags in (6.5); they are available from April 1974 to November 1996.

The source for the data on the dollar value of exports is the U.S. Bureau of the Census. To convert nominal exports into real (domestic output) exports x_t we use the U.S. wholesale price index from the OECD Main Economic Indicators. This is also the source for foreign industrial production, which is commonly used to proxy y_t^*, because real national income is only available at the quarterly frequency. The monthly nominal exchange rate is taken from the IMF International Financial Statistics and the OECD wholesale price indices are used to convert it to the real exchange rate s_t (except for the French real exchange rate, which is based on French and U.S. consumer price indices, because French WPI is not available).

To obtain a measure for $E_{t-l}\{s_t\}$ we simply take the lagged rate s_{t-l}. For short horizons such a random walk point forecast outperforms forecasts from structural exchange rate models (see Meese and Rogoff (1983)). For long horizons, however, there appears to be some predictability in real exchange rate changes using fundamentals. Nevertheless, a random walk based forecasting rule is a good approximation (see Meese and Rogoff (1983), and Mark and Choi (1997) for empirical evidence).

Measuring exchange rate risk, $V_{t-l}^{1/2}\{s_t\}$, is less obvious. Given the importance of this variable, we discuss it extensively in the next subsection. Our preferred measure will appear an AR(2) based forecast using past monthly real exchange rate volatilities, where monthly volatility is defined as the square root of the sum of squared daily percentage changes in that month.[8]

[8] Although daily nominal exchange rates are observable (from Datastream), daily real exchange rates are not perfectly observable, because price ratios P_t^*/P_t are only observed once every month. However, given the stability of the price ratios, we use good proxies of daily price ratios by linear interpolation of the monthly ratios.

6.3.2 Real Exchange Rate Risk Measure

In this subsection we first discuss and compare two risk measures that are commonly used in the trade literature. Then we introduce an alternative measure, based on daily exchange rates, which provides a more appropriate description of risk. Two characteristics of this risk measure will play a crucial role in the derivation of the conclusion of the chapter.

The measures used in the trade literature so far are typically one-period-ahead volatility measures, that is, $V_{t-1}^{1/2}\{s_t\}$ instead of $V_{t-l}^{1/2}\{s_t\}$ for some positive l. Hence, in case of monthly data it is one-month-ahead risk and for quarterly data it is one-quarter-ahead risk that is allowed to affect trade flows. Although one should not a priori impose a specific time lag, for ease of exposition we first discuss the various risk terms for one-period-ahead risk. After that, we come back to multi-period-ahead risk and explain how we quantify it.

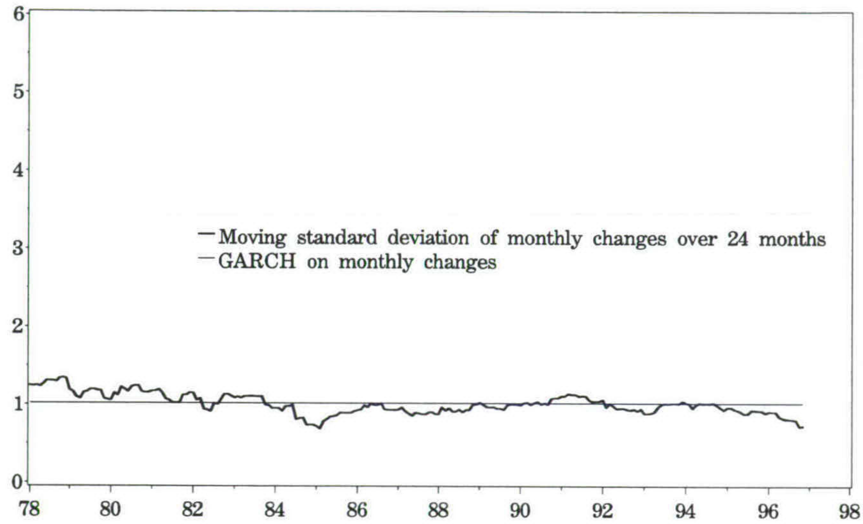
The first commonly used risk measure is the moving sample standard deviation of past percentage real exchange rate changes. The window width is prespecified and is usually about two years (for instance, Chowdury (1993) uses eight quarters). For illustrative purposes, let us therefore assume that the window width is 24 months, so that the moving standard deviation measure becomes

$$V_{t-1}^{1/2}\{s_t\} = \sqrt{\frac{1}{24}\sum_{m=1}^{24}[100(s_{t-m} - s_{t-m-1})]^2}, \tag{6.6}$$

where it is implicitly assumed that the average real exchange rate change is zero. One can interpret measure (6.6) as first approximating volatility in month t by $[100(s_t - s_{t-1})]^2$ and then smoothing by taking the average over 24 months. Of course, taking a 24-months equally-weighted average is rather ad hoc, but usually the authors report that the results are not very sensitive to other weighting schemes (see Chowdury (1993), among others).

The main characteristic of the moving standard deviation measure (6.6) is that it implies a high (24 months) persistence of real exchange rate shocks and, therefore, considerable serial correlation in risk. This is illustrated by figures 6.1A and 6.2A, in which the thick lines plot measure (6.6) for the two most important trading partners of the U.S., namely Canada and Japan, respectively. Apart from the monthly shocks, there are some long swings in the risk measure, particularly for Japan. Later on in this subsection we will check whether the high autocorrelation is real or spuriously induced by definition (6.6).

A: Common risk measures: using monthly data up to one month ago



B: Alternative risk measure: AR(2) based on sum of squared daily changes in a month

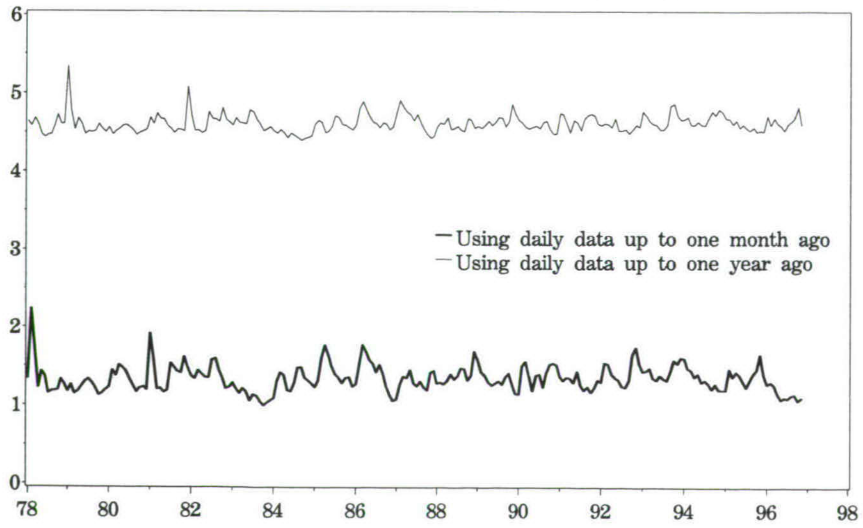


Figure 6.1: Risk measures for Canadian dollar real exchange rate

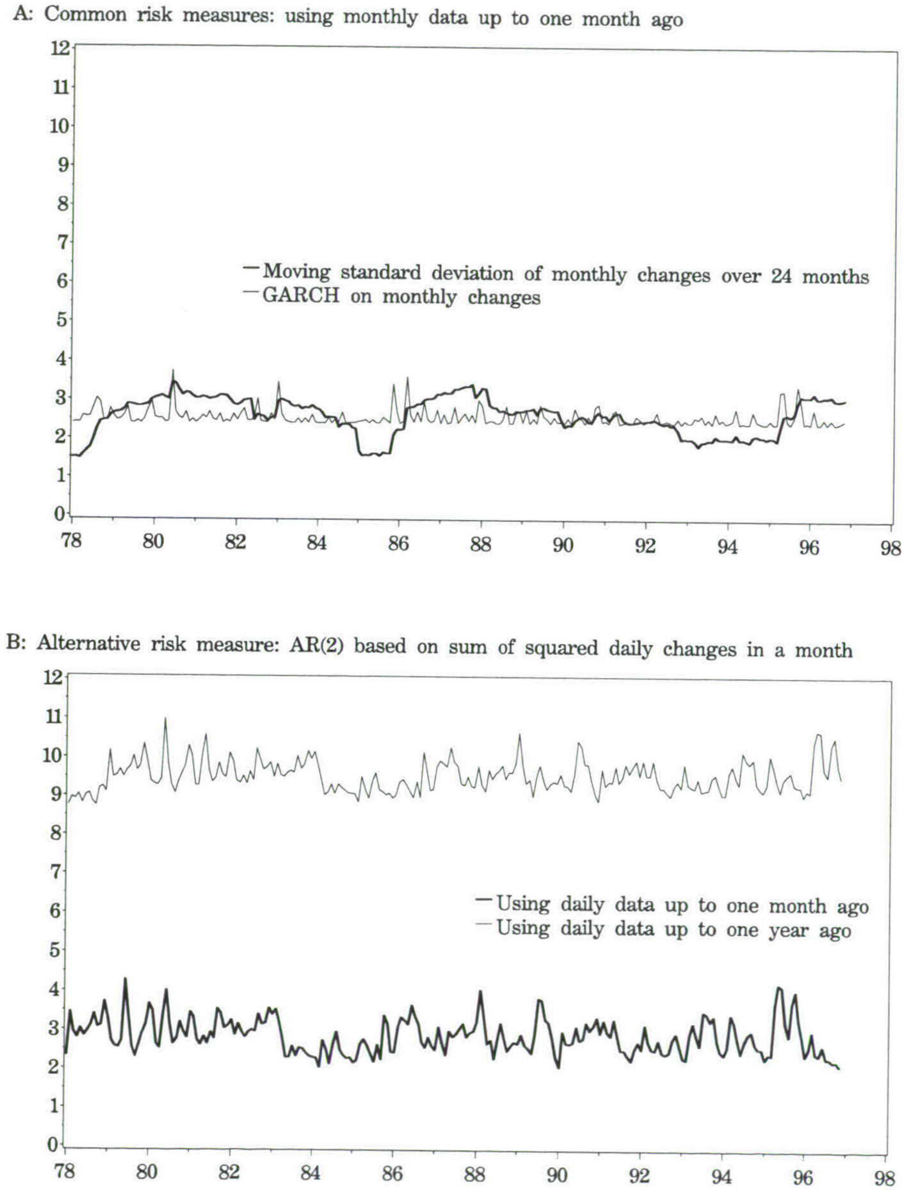


Figure 6.2: Risk measures for Japanese yen real exchange rate

The second measure of exchange rate risk that is commonly used in the trade literature is based on a GARCH model to smooth monthly volatilities $[100(s_t - s_{t-1})]^2$. For instance, if one uses a GARCH(1,1) model, the risk measure is

$$V_{t-1}^{1/2}\{s_t\} = \sqrt{\omega_0 + \omega_1[100(s_{t-1} - s_{t-2})]^2 + \omega_2 V_{t-2}\{s_{t-1}\}}, \tag{6.7}$$

where we assume for the surprise term $[100(s_{t-1} - s_{t-2})]^2$ that the mean real exchange rate change is zero.

The main characteristic of measure (6.7) regarding our purpose of measuring volatility at the monthly frequency is illustrated by the thin lines in figures 6.1A and 6.2A. They show that there is low persistence of shocks in risk; for Canada the GARCH approach even results in constant risk. The reason for this becomes clear from table 6.1. The top half of that table presents the first-order autocorrelation, ρ_1, and the Box-Pierce combination Q_{10} of the first ten autocorrelations of monthly volatility $\sqrt{[100(s_t - s_{t-1})]^2}$. It demonstrates that squared real exchange rate changes exhibit zero or small autocorrelation at the monthly frequency (we always use a significance level of 5%). This result is well-known from the GARCH literature (see Bollerslev, Chou and Kroner (1992)) and causes the low or zero autocorrelation in the monthly GARCH risk measures.

The low serial correlation in risk found by measure (6.7) is not consistent with the high correlation suggested above by the moving standard deviation measure (6.6). Hence, what is the true degree of serial correlation?

To analyze this question we start from an idea presented by Merton (1980) and formalized by Andersen and Bollerslev (1998). The latter authors argue that the ex-post squared change in a period is a very noisy indicator for the latent variance in that period. They propose to measure volatility by cumulating squared high-frequency changes in the period, so as to decrease measurement error. Under the reasonable assumption of no autocorrelation in the high-frequency changes, they argue that, as the observation frequency tends to infinity, the cumulative measure converges to the true volatility.

We use this idea to reduce the noise in the monthly volatilities $[100(s_t - s_{t-1})]^2$. That is, we measure monthly volatility by the sum of squared daily percentage real exchange rate changes over all days in that month, $\sqrt{\sum_{d \in D_t}[100(s_d - s_{d-1})]^2}$, where D_t is the set of days in month t (see also Merton (1980) on stock returns). As each monthly volatility is now based on about 21 daily volatilities, it is not

Table 6.1: Autocorrelation in monthly real exchange rate volatility

		U.S. dollar real exchange rate versus currency of					
		Can	Fra	Ger	Ita	Jap	U.K.
Using monthly data i.e. $\sqrt{[100 \cdot (s_t - s_{t-1})]^2}$	ρ_1	0.06 (0.06)	0.00 (0.06)	-0.01 (0.06)	0.21* (0.06)	0.12* (0.06)	0.15 * (0.06)
	Q_{10}	15.07 [0.13]	10.68 [0.38]	14.73 [0.14]	52.82* [0.00]	17.11 [0.07]	21.42 * [0.02]
Using daily data i.e. $\sqrt{\sum_{d \in D_t}[100(s_d - s_{d-1})]^2}$	ρ_1	0.27* (0.06)	0.48* (0.05)	0.48* (0.05)	0.50* (0.05)	0.43* (0.05)	0.55 * (0.05)
	Q_{10}	48.14* [0.00]	255.19* [0.00]	215.76* [0.00]	243.81* [0.00]	253.43* [0.00]	320.22 * [0.00]

Standard errors in parentheses and p-values in square brackets; * is significant at 5% level.
The symbol ρ_1 denotes the first-order autocorrelation and Q_{10} is the Box-Pierce statistic of order 10.

surprising that this measure is more accurate than the monthly volatility measure based on a single monthly change.

We now re-examine the serial correlation in monthly volatility with the new measure. The second half of table 6.1 shows that there is clear evidence of serial correlation. This indicates that our GARCH based claim of no or low autocorrelation is wrong, a result previously documented by Andersen, Bollerslev, Diebold and Labys (1999).

To analyze whether the serial correlation in volatility is high, as the moving standard deviation measure (6.6) suggests, we estimate an autoregressive model for the monthly volatilities (based on daily data). As table 6.2 demonstrates, AR(2) models with moderate AR coefficients suffice to capture all serial correlation. Hence, the suggestion of high persistence of shocks from the moving standard deviation measure is not correct either. We conclude that there is significant autocorrelation in monthly volatilities, but that it dies out rather quickly.

Given the drawbacks of the moving standard deviation and GARCH measure for our purpose of studying the effect of exchange rate risk on trade, we propose an alternative risk measure. It is based on the AR(2) estimates just presented. More specifically, our measure is the AR(2) forecast based on past monthly volatilities obtained from daily data, that is,

$$V_{t-1}^{1/2}\{s_t\} = \mu_v + \sum_{p=1}^{2} \alpha_p\Big(\sqrt{\sum_{d \in D_{t-p}} [100(s_d - s_{d-1})]^2} - \mu_v\Big), \tag{6.8}$$

Table 6.2: AR(2) estimation results for monthly real exchange rate volatility

		U.S. dollar real exchange rate versus currency of					
		Can	Fra	Ger	Ita	Jap	U.K.
Mean	μ_v	1.32 (0.05)	2.71 (0.14)	2.85 (0.13)	2.68 (0.17)	2.72 (0.13)	2.68 (0.16)
AR coefficients	α_1	0.23* (0.06)	0.41* (0.06)	0.44* (0.06)	0.41* (0.06)	0.35* (0.06)	0.43 * (0.06)
	α_2	0.14* (0.06)	0.16* (0.06)	0.10 (0.06)	0.18* (0.06)	0.18* (0.06)	0.23 * (0.06)
Residual diagnostics							
Autocorrelation	ρ_1	-0.01 (0.06)	-0.01 (0.06)	-0.01 (0.06)	-0.01 (0.06)	-0.02 (0.06)	-0.01 (0.06)
	Q_{10}	9.34 [0.50]	15.89 [0.10]	10.71 [0.38]	14.22 [0.16]	15.34 [0.12]	3.44 [0.97]

Standard errors in parentheses and p-values in square brackets; * is significant at 5% level.
Definitions of ρ_1 and Q_{10}: see notes of table 6.1.

where μ_v, α_1 and α_2 are substituted by the estimates presented in table 6.2. Because this measure takes account of the serial correlation in monthly volatilities in a better way than the two commonly used risk measures described above, thereby reducing measurement error for the important exchange risk variable, we use it in the remaining part of the chapter.

An additional advantage of our measure is that multi-month-ahead risk, $V_{t-l}^{1/2}\{s_t\}$ for some positive l, which is the measure we actually need in (6.5), is easy to compute. Assuming that monthly real exchange rate changes are uncorrelated, $V_{t-l}^{1/2}\{s_t\}$ is the square root of $V_{t-l}\{s_{t-l+1}\} + V_{t-l}\{s_{t-l+2} - s_{t-l+1}\} + \ldots + V_{t-l}\{s_t - s_{t-1}\}$, where each term is a standard multi-period-ahead AR(2) forecast, which can be obtained in a recursive manner.

Two characteristics of (the multi-month-ahead version of) risk measure (6.8) will play a crucial role in subsection 6.5.2, where we derive the final conclusion of the chapter. These characteristics concern the variation in risk over time and the duration of deviations from average risk. Figures 6.1B and 6.2B illustrate the risk measure for Canada and Japan, respectively, for both $l=1$ and $l=12$. They show that real exchange rate risk is time-varying, but that shocks do not persist very long in risk. Moreover, particularly for $V_{t-12}\{s_t\}$, the time-variation in risk is small relative to the risk level. This conclusion is supported by table 6.3, as

Table 6.3: Descriptive statistics of real exchange rate risk measure

		U.S. dollar real exchange rate versus currency of					
		Can	Fra	Ger	Ita	Jap	U.K.
$V_{t-1}^{1/2}\{s_t\}$	mean	1.35	2.83	2.94	2.67	2.85	2.79
	coeff. of variation	0.13	0.20	0.18	0.17	0.16	0.22
$V_{t-12}^{1/2}\{s_t\}$	mean	4.61	9.49	9.95	9.29	9.53	9.42
	coeff. of variation	0.02	0.06	0.04	0.05	0.04	0.07

The risk measure $V_{t-l}^{1/2}\{s_t\}$ $(l = 1, 12)$ is the l-months-ahead forecast based on the AR(2) process that has been estimated for the monthly real exchange rate volatilities from daily data. See the discussion below (6.8) for an exact description.

the standard deviation of risk is on average only 5% of the mean.

6.3.3 Non-Stationarity and Cointegration

To specify a time series model for exports in section 6.4 using the four variables of economic model (6.5), we first have to investigate the stationarity of these variables. It is common to assume that two of these, real exports x_t and foreign industrial production y_t^*, have a unit root. In contrast, measure (6.8) for exchange rate risk $V_{t-l}^{1/2}\{s_t\}$ is stationary, as the AR(2) estimates in table 6.2 are positive and their sum is well below unity (see Hamilton (1994, p. 57)). Stationarity is confirmed by plots of the risk measure; see figures 6.1B and 6.2B for Canada and Japan, respectively. Finally, we assume that the expected real exchange rate, $E_{t-l}\{s_t\} = s_{t-l}$, is stationary. This is based on the recent literature on purchasing power parity (PPP), which provides more and more evidence of long-run relative PPP, in other words, of stationarity of the real exchange rate (for instance, see Abuaf and Jorion (1990), Koedijk, Schotman and Van Dijk (1998) and Chapter 4).[9]

Next, we check for cointegration between the two unit root variables x_t (exports) and y_t^* (foreign industrial production). From an economic point of view one expects that they are cointegrated. This is confirmed by the empirical results in Sawyer and Sprinkle (1997), among others. But obtaining statistical evidence

[9]If one is not willing to assume stationarity of the real exchange rate, the main conclusion of the chapter, which concerns the stationary risk measure, is still valid; this follows from subsection 6.5.1.

for our data is not so obvious, as augmented Dickey-Fuller unit root tests (not reported) on the residuals from a regression of x_t on y_t^* do not show evidence of cointegration.

The insignificant Dickey-Fuller test results, however, do not imply the absence of cointegration, as it is well-known that standard unit root tests may have problems with power. To examine this, we inspect the residual plots concerning the regression of x_t on y_t^*. They demonstrate that there is no trend in the residuals and that the residuals exhibit long swings. For instance, for all six flows the residuals swing downwards for some years before 1986 and follow an upward swing in the years after that. These long swings, taking several years, in combination with the moderate length of our export series (19 years) may well be the reason for the insignificant Dickey-Fuller tests. After all, the swings in the residual series have a similar shape as those in the real exchange rates (which are likely to be the cause of the residual swings), and from the PPP literature we know that standard unit root tests have great difficulties in finding stationarity from short stationary series exhibiting long swings.

Although economic intuition argues for cointegration, we still have no conclusive statistical evidence. Obtaining such evidence requires a much more detailed cointegration analysis, which goes beyond the scope of this chapter. Instead, we follow an indirect approach. First, we simply assume cointegration and specify the econometric model using x_t and y_t^* in levels. Afterwards, having estimated the model, we examine the residuals of that model. We will show in subsection 6.5.1 that they are stationary, so that, given the stationarity of $E_{t-l}\{s_t\}$ and $V_{t-l}^{1/2}\{s_t\}$, it is very likely that x_t and y_t^* are cointegrated, as economic intuition suggests.

6.4 Econometric Model

In this section we develop the econometric model to be estimated later on. Its main elements are the export equation, described in subsection 6.4.1, and the restrictions placed on its dynamic structure, discussed in 6.4.2.

6.4.1 Export Equation

To specify an econometric equation for real U.S. exports we use the variables that appear in economic model (6.5). That model takes explicit account of the important dynamic nature of international trade by specifying the determinants of exports in month t when the export contract was signed l months before. However, the data on U.S. exports are aggregated across all products and it is likely that for different products the lags l are different. To account for this, we use a distributed lag model, where the effect of a change in a regressor is allowed to be distributed over time.

Given the assumed cointegration between real exports x_t and foreign industrial production y_t^*, the stationarity of $E_{t-l}\{s_t\}$ and $V_{t-l}^{1/2}\{s_t\}$, and assuming linearity, we specify real exports as

$$x_t = \beta_0 + \sum_{l=1}^{\infty} \left(\beta_{yl} y_{t-l}^* + \beta_{El} E_{t-l}\{s_t\} + \beta_{Vl} V_{t-l}^{1/2}\{s_t\}\right) + \varepsilon_t, \tag{6.9}$$

where the disturbance term ε_t is allowed to follow an AR(2) process with autoregressive coefficients θ_1 and θ_2 and with conditionally normal innovations having variance σ^2.[10] Although x_t concerns bilateral exports, we suppress the index indicating the partner country for notational simplicity. We also do not explicitly write down the eleven monthly dummies that we include to correct for seasonal effects.

Of course, unrestricted estimation of (6.9) is not feasible because of the infinite number of parameters. In the next subsection we introduce the restrictions on β_{yl}, β_{El} and β_{Vl} that complete the econometric model.

6.4.2 Poisson Lag Structure

Careful investigation of the lag structure is important for dynamic trade equations such as (6.9). This subsection pays special attention to the lags. We first discuss two popular lag structures. After that, we introduce an alternative structure based on the Poisson probability (mass) function, which we argue is more appropriate. Moreover, the Poisson lag structure allows us to let the data reveal

[10]For Canada we allow for a break in β_y from 1991 onwards to account for the increase in trade openness due to the Free Trade Agreement between the U.S. and Canada. Moreover, we use an AR(5) instead of AR(2) process to capture all autocorrelation in the disturbance term.

the exchange risk horizon that is relevant for goods traders, which is an important element in the derivation of the main conclusion of the chapter in subsection 6.5.2.

In the literature there exist several ways of restricting the infinite number of coefficients $\beta_1, \beta_2, \ldots$ in (6.9) to obtain a parsimonious model (β_l is shorthand notation for β_{yl}, β_{El} or β_{Vl}). For instance, one can use a geometric lag specification, that is, $\beta_l = \beta \cdot w_l$, where $w_l = \gamma \cdot (1-\gamma)^{l-1}$ is the geometric probability function translated one unit to the right ($0 < \gamma < 1$). It implies that the β_l are decreasing over l. This may be appropriate for the income effects β_{yl}, as there appears to be some agreement in the literature that income effects are large for small lags and decline rapidly thereafter (see Goldstein and Khan (1985)). However, according to Goldstein and Khan there is much less of a consensus on the lag pattern for the expected exchange rate effects β_{El}; that may well be hump shaped, as Sawyer and Sprinkle (1997) claim. Hence, it is not appropriate to impose a geometric lag specification a priori.

A second example of a popular lag structure is the polynomial or Almon lag specification. It assumes that the β_l fall on a polynomial of a prespecified order. Such a specification is more flexible with respect to the dynamics of β_l than the geometric model, as it allows for both a declining and a hump shaped lag pattern. However, it may well occur that the polynomial structure forces some β_l to be positive and others to be negative. This is difficult to justify theoretically (see Goldstein and Khan (1985)).

Given the importance of a satisfactory lag structure, we introduce an alternative approach to avoid the problems just described. Let us suppose that all β_l have the same sign. Then, one can write $\beta_l = \beta \cdot w_l$, where $w_l \geq 0$ and $\sum_{l=1}^{\infty} w_l = 1$. Hence, β gives the total, long run effect of the regressor. The w_l, on the other hand, describe how the total effect is distributed over time; by definition, they form a probability function with support $\{1, 2, \ldots\}$.

Besides the convenient interpretation of β and the w_l, the main attractive feature of our class of probability function based lag specifications is its flexibility. One can choose any probability function for the w_l, depending on the specific needs. For instance, the approach encompasses the geometric lag specification as the special case where the w_l are defined by a translated geometric probability function (see above). It can also capture, for instance, hump shaped or bimodal lag patterns.

Within the class of lag specifications just described, we take "Poisson lags" for our export model (6.9). That is,

$$\beta_{il} = \beta_i \cdot \frac{(\lambda_i - 1)^{l-1}}{(l-1)!} \exp[-(\lambda_i - 1)], \qquad \text{for } \lambda_i \geq 1 \text{ and } i = y, E, V. \tag{6.10}$$

Note that we have to translate the Poisson probability function one unit to the right, because l starts at one instead of zero. The parameter λ is close to the mode of the translated Poisson distribution.[11] Hence, we give λ the convenient interpretation of the lag at which the maximal effect occurs, that is, the lag with the largest coefficient β_l. Because λ_E and λ_V both concern the exchange rate distribution (mean and variance) and to avoid identification problems if β_E or β_V is zero, we impose that λ_E and λ_V are equal to, say, λ_{EV} (this restriction will be tested in subsection 6.5.1). We allow λ_y and λ_{EV} to be different.

The Poisson lag structure (6.10) is very parsimonious. This is at the cost of flexibility. However, Poisson lags can capture a declining lag structure as well as a hump shaped one and imply that all β_l have the same sign. Hence, Poisson lags avoid the disadvantages concerning geometric lags and polynomial lags discussed above. We can let the data decide whether a declining or hump shaped lag structure is more appropriate and how long it takes for industrial production and exchange rates to have the strongest effect on exports, an issue that is also unresolved in the literature (see Sawyer and Sprinkle (1997)). Figure 6.3 illustrates the Poisson lags for $\lambda = 3.38$ and $\lambda = 12.85$ (with $\beta = 2.23$ and $\beta = 0.62$, respectively; the numbers are based on the estimation results to be discussed below).

This completes the description of the econometric model for the determination of exports. It is given by (6.9) and (6.10).

6.5 Empirical Results

In this section we first report the estimates of the parameters in the model just developed. As in the existing literature, we find an ambiguous effect of exchange

[11] The exact mode of the translated Poisson distribution with parameter λ is the largest integer l less than λ; if λ itself is an integer, then $l = \lambda - 1$ and $l = \lambda$ are tie modes.

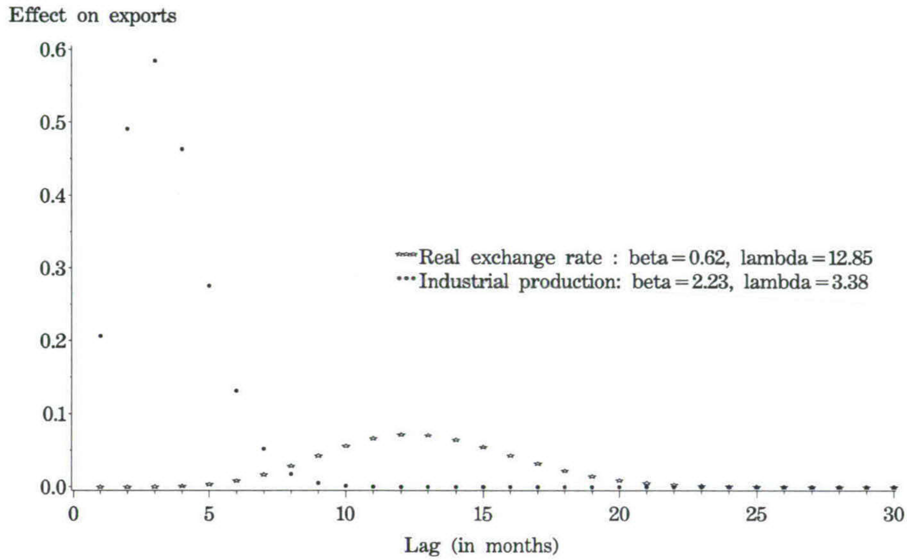


Figure 6.3: Distribution of total effect β of regressors on exports over time according to a Poisson(λ) lag structure

rate risk on exports. In subsection 6.5.2 we provide an explanation for that.

6.5.1 Estimation Results

We estimate the econometric model of section 6.4 with maximum likelihood (ML) on each of the six U.S. export flows separately.[12] Table 6.4 present the results.

The focus parameter of this chapter is β_V, the total impact of real exchange rate risk on exports. We find that the estimate of β_V is significantly positive for Canada, significantly negative for Italy and insignificant for the other four countries. Hence, as in the existing literature, we find no clear effect of risk on

[12]Multivariate ML is theoretically possible. However, the cross-sectional correlation in the univariate residuals turns out to be low (the average absolute correlation between the residuals of two equations is only 0.12, and the maximum is 0.25), so that the efficiency gains from multivariate estimation are likely to be small. Moreover, multivariate estimation involves more than one hundred parameters, so that there is a serious danger of ending up in a local maximum of the likelihood function.

Table 6.4: Estimation results for export equations

		U.S. exports to					
		Can	Fra	Ger	Ita	Jap	U.K.
Constant	β_0	2.40 (1.55)	-7.01* (0.82)	2.11* (0.42)	6.54* (0.67)	5.43* (0.41)	1.71 * (0.49)
Foreign industr. prod.	β_y	1.80* (0.07)	4.18* (0.09)	2.25* (0.04)	1.38* (0.17)	1.28* (0.09)	2.49 * (0.07)
Lag of max. effect	λ_y	2.48 (0.44)	5.63 (0.86)	2.87 (0.63)	2.51 (0.73)	3.41 (1.69)	11.05 (1.54)
Expected exch. rate	β_E	0.50* (0.18)	0.50* (0.08)	0.65* (0.05)	0.52* (0.09)	0.95* (0.10)	0.62 * (0.09)
Exchange rate risk	β_V	0.62* (0.20)	0.05 (0.03)	0.01 (0.03)	-0.08* (0.03)	0.04 (0.04)	-0.04 (0.03)
Lag of max. effect	λ_{EV}	17.61 (2.19)	10.34 (1.78)	9.91 (1.15)	8.50 (1.19)	12.97 (1.05)	17.77 (1.96)
AR(2) for error	θ_1	0.25* (0.06)	0.27* (0.07)	0.24* (0.07)	0.37* (0.07)	0.45* (0.08)	0.26 * (0.06)
	θ_2	0.03 (0.08)	0.30* (0.08)	0.14 (0.07)	0.17* (0.07)	0.23* (0.07)	0.19 * (0.07)
Error variance	$100\sigma^2$	0.29 (0.03)	0.65 (0.06)	0.50 (0.05)	0.81 (0.09)	0.36 (0.04)	0.82 (0.09)
Log-likelihood		342.04	249.75	280.14	223.93	318.19	223.63
Residual diagnostics							
Autocorrelation	ρ_1	0.01 (0.07)	-0.04 (0.07)	-0.01 (0.07)	-0.02 (0.07)	-0.02 (0.07)	-0.02 (0.07)
	Q_{10}	2.31 [0.99]	7.57 [0.67]	11.89 [0.29]	11.80 [0.30]	7.97 [0.63]	12.15 [0.28]
Autocorr. squares	ρ_1^s	0.13 (0.07)	0.03 (0.07)	0.03 (0.07)	-0.00 (0.07)	-0.10 (0.07)	0.11 (0.08)
	Q_{10}^s	15.29 [0.12]	5.91 [0.82]	5.17 [0.88]	16.27 [0.09]	17.72 [0.06]	7.61 [0.67]

Standard errors in parentheses and p-values in square brackets; * is significantly different from zero at 5% level.

The estimated equation is (6.9) with the Poisson lag restriction (6.10); we do not report the estimates for the monthly seasonality dummies.

The significance of the estimates for β_y is based on the cointegration between x_t and y_t^*. The significance of the estimates for β_E is based on t-ratios. Because of the slow mean reversion in real exchange rates, the asymptotic 5% critical value of about two is possibly different from the critical value relevant for our finite sample. Nevertheless, we consider the t-ratios to be sufficiently large to conclude that the estimates are significant.

For exports to Canada we have allowed for a break in β_y from 1991 onwards to account for the increase in trade openness due to the Free Trade Agreement between the U.S. and Canada; the estimated increase in β_y is 0.04* (0.005). Moreover, we have estimated an AR(5) instead of AR(2) process to capture all autocorrelation in the error term; the three extra AR parameter estimates are 0.26* (0.08), -0.14* (0.07) and 0.17* (0.07).

Definitions of ρ_1 and Q_{10}: see notes of table 6.1; ρ_1^s and Q_{10}^s are similarly defined, except that they concern the squared residuals.

exports.

Table 6.4 also demonstrates that foreign industrial production has the expected positive effect on the real (domestic output) value of U.S. exports. This holds for all six series. The average estimate of β_y is 2.23.[13]

An attractive implication of the Poisson lag specification (6.10) is that we can directly estimate the time lag λ_y between a change in industrial production and the maximal change in exports. Table 6.4 shows that the maximal effect occurs after about one quarter (the average estimate of λ_y is 3.38, ignoring the outlying estimate for the U.K.). This conclusion is robust to the use of another lag specification, as a preliminary analysis with polynomial lag structures of various degrees points in the same direction. Hence, the effect of foreign income on U.S. exports goes quite rapidly; this corroborates Goldstein and Khan (1985) and Sawyer and Sprinkle (1997). The dots in figure 6.3 illustrate the implication of the average λ_y for the distribution of the average β_y over the lags.

The remaining regressor is the expected real exchange rate. As table 6.4 demonstrates, all six estimates for β_E are significantly positive. This is not surprising, as a U.S. dollar depreciation generally lowers the foreign currency price of (dollar denominated) U.S. exports, thereby increasing the quantity and dollar value of exports. The average estimate of β_E is 0.62. It is remarkable that the values of our estimates are so consistent across countries given the wide range of estimated export price elasticities in the literature, as analyzed by Marquez (1999). This consistency is a sign of robustness of our model.

From the Poisson lag structure we find that the maximal effect of the exchange rate distribution occurs after about one year (the average λ_{EV} is 12.85).[14] This conclusion is again supported by a preliminary analysis with polynomial lags of various orders. Therefore, the short-run effect of changes in the exchange rate distribution on exports is small, while in the longer run there is a clear effect. This

[13]The estimates for β_y are not directly comparable with the income elasticities of U.S. exports that are typically reported in the literature, since the endogenous variable in (6.4.1) is the value of exports, not the quantity, and because the explanatory variable is industrial production, not real national income.

[14]Recall that λ_{EV} determines the lag distribution of both β_E and the risk coefficient β_V (see assumption $\lambda_E = \lambda_V = \lambda_{EV}$ below (6.10)). To test the restrictiveness of that assumption we perform a likelihood ratio test. The likelihood ratios [p-value] are 0.91 [0.34] for Canada, 3.80 [0.05] for France, 0.36 [0.55] for Germany, 0.01 [0.92] for Italy, 3.62 [0.06] for Japan, and 0.55 [0.46] for the U.K.. Hence, we do not reject the restriction.

supports the view of a hump shaped instead of a declining lag pattern and hence helps solve the question on the true lag pattern for exchange rates (Goldstein and Khan (1985)). The stars in figure 6.3 illustrate the distribution of the average β_E over the lags as implied by the average λ_{EV}.

The final estimation results presented in table 6.4 concern the autoregressive parameters of the AR process for the error term ε_t in (6.9). The moderate values for the estimates of θ_1 and θ_2 show that the systematic part of export equation (6.9) describes the dynamics of exports quite well. Moreover, the fact that the estimates of θ_1 and θ_2 are positive and that their sum is well below unity ensures that the estimated AR process is stationary (see Hamilton (1994, p. 57)). Stationarity is also confirmed by the residual plots (not presented). This supports our assumption of cointegration between the trending variables exports and industrial production, as made in subsection 6.3.3.

Table 6.4 also reports some diagnostic statistics. There is no sign of remaining autocorrelation or conditional heteroskedasticity in the residuals, so that we have no reason to extend the model.

6.5.2 Why is the Effect of Exchange Risk on Exports Ambiguous?

As just discussed, we find no clear evidence of an effect of real exchange rate risk on the real (domestic output) value of exports. In this subsection we try to explain this.

We distinguish two points of view. First, it may be that there is no effect of risk on trade; this would imply that the common idea of a negative effect is wrong. Second, there is an effect, but it is overshadowed by the variation in the unsystematic part of the model in such a way that one cannot discover the true effect of risk on trade from the limited time series that are typically available.

In the literature there is a tendency towards the first point of view, because the many studies on this issue have not yet come to a conclusive answer. We, however, argue that the second point may be more relevant.

This claim is based on the estimated Poisson parameter λ_{EV} (the lag with the maximal exchange rate effect on exports) and on the two main characteristics of real exchange rate risk as discussed at the end of subsection 6.3.2. From the estimated λ_{EV} we concluded that the maximal effect of exchange rates on trade occurs after about one year. We have seen that, at such a long horizon, the variation of exchange risk over time is rather small (see table 6.3 and figures 6.1B and 6.2B, particularly the one-year-ahead risk measure). Moreover, the second characteristic of risk discussed in subsection 6.3.2 shows that deviations from average risk are short-lived, since AR(2) processes with moderate autoregressive parameters are already sufficient to capture the autocorrelation in risk (see table 6.2 and figures 6.1B and 6.2B).

The three properties imply that, even if risk affects exports, the effect explains only little of the variation and the long-term movements in exports over time; other shocks to exports are likely to dominate and overshadow such an effect. Loosely speaking, risk is too constant to identify its effect on exports from time series analysis. We conclude it is unlikely that one will discover the true effect of risk on exports from the limited time series data that are typically available, no matter whether the true effect is zero or not.

6.6 Conclusion

This chapter presents an empirical study on monthly bilateral aggregate U.S. exports to the other G7 countries from 1978 to 1996. To motivate the choice of variables in the econometric model we develop an economic model, where we explicitly account for the time lag between the export decision and the actual trade flow and payment. The model implies that not only foreign income and the expected future real exchange rate are important, but also that real exchange rate risk may be relevant for exports. This latter effect is the main focus of the chapter. In particular, why are its empirical estimates in the literature so ambiguous, even though most economists think that the effect is negative?

From a methodological point of view, the chapter yields two contributions to the trade literature. First, we improve on currently used risk measures by using daily exchange rates to construct multi-month-ahead risk. This reduces measurement error and makes the estimated effect of risk on exports more accurate. In addition, we pay special attention to the dynamic structure of the model by introducing a convenient Poisson lag structure for the distributed lag model.

The empirical results demonstrate that, as expected, foreign income affects

U.S. exports positively and rather quickly, since the maximal effect in the Poisson lag structure occurs after about one quarter. Exports react much slower to changes in the real exchange rate distribution, as the maximal effect happens only after about one year. The expected real exchange rate level has the normal positive effect, but real exchange rate risk has no clear effect.

To explain this latter, commonly reported finding, we examine the long-term (about one year) risk that is relevant for goods traders in more detail. Such long-term risk appears rather constant over time with only short-term deviations from average risk. In our opinion, this is the reason why it is so difficult to find an effect of exchange rate risk on trade from time series data.

It is important to realize that our conclusion concerns countries with low time-variation in long-term real exchange rate risk, such as most developed countries over the post Bretton Woods period. It would be interesting to study the effect of risk on trade flows between countries with more time-variation in risk, for instance, developing countries. In addition, employing cross-sectional variation in exchange risk may be fruitful. Such panel or pure cross-sectional studies may benefit from the few cross-sectional papers that already exist and that tend to be more supportive for a negative effect of exchange risk on trade (see Côté (1994)). This is left for future research.

Chapter 7

Conclusion

This thesis contains empirical essays on exchange rates and their effects on international goods trade. Its contribution to the literature is twofold. First, we provide an empirical answer to a number of significant economic questions, using the United States as the central country of analysis. Second, we refine existing econometric methods or, if necessary, introduce new techniques to analyze the economic question at hand. In this concluding chapter we briefly summarize the most important results per essay and relate them to each other; this will explain the contribution of this thesis in more detail. We also give suggestions for future research.

Chapter 2: Improving GARCH Volatility Forecasts with a New Regime-Switching GARCH Model

Chapter 2 concerns forecasting the volatility of exchange rates. The GARCH model is often used quite successfully in this respect. However, GARCH volatility forecasts appear too variable. This is likely to be caused by the high estimated persistence of shocks in the GARCH forecast. To obtain more flexibility regarding volatility persistence, we generalize the GARCH model to regime-switching GARCH. Shocks can now also be "pressure relieving" instead of persistent.

Our specification of regime-switching GARCH is novel. The most important difference with the existing version of Gray (1996a) is that our specification yields a much more convenient volatility forecasting formula. This allows us to compare regime-switching GARCH with standard GARCH forecasts.

The empirical application concerning the three main U.S. dollar exchange rates (dollar versus German mark, Japanese yen and U.K. pound) shows that the regime-switching GARCH (out-of-sample) forecasts are no longer too variable and that they substantially outperform the standard GARCH forecasts.

The theoretical and empirical improvements just described make our regime-switching GARCH specification potentially useful for future research. For example, the proposed technique of averaging out the unobserved volatility regime may also be applied in other regime-switching models, such as the regime-switching mean models in Chapters 3 and 4. Moreover, the empirical enhancements may make regime-switching GARCH useful when analyzing the volatility of other financial variables, such as stock returns, or when studying the effect of volatility on, say, option pricing.

Chapter 3: Long Swings in Exchange Rates: Are They Really in the Data?

Chapter 3 tests the often-used random walk model for exchange rates against a model that allows for long swings. The latter model is again a regime-switching model, but not in the sense of Chapter 2. We now allow for switches in the mean exchange rate depreciation, while Chapter 2 models switches in the variance. The long swings, or regime-switching model seems more plausible than the random walk from an economic point of view, because it allows for the fact that changes in economic policy can cause switches in the exchange rate regime.

Quarterly data on the three main dollar exchange rates do not yield evidence of long swings. However, weekly data on the same series reveal significant evidence. Apparently, the sampling frequency can matter for tests on long run phenomena such as long swings. Hence, we conclude that long swings are in the data, but that finding them requires fairly high frequency data.

In addition, we find that taking account of long swings in forecasting helps predict the direction of exchange rate changes.

Both empirical results suggest that exchange rates do not behave in a purely random way and that there are occasional changes in regime. Hence, research on the probability of such changes seems promising. For instance, one can directly analyze the role of policy changes for switches in the exchange rate regime. In this respect, monetary policy announcements may have an effect, as in Kaminsky (1993). Moreover, one can test whether foreign exchange interventions affect the exchange rate regime-switching probabilities, as such interventions may signal changes in future monetary policy (see Loopesko (1984)). One can also include market fundamentals in the regime-switching probabilities, such as the trade balance disequilibrium and deviations from purchasing power parity. (in Chapter 4 we examine the role of deviations from purchasing power parity). The empirical results can then be used to improve existing theoretical exchange rate models.

Chapter 4: Purchasing Power Parity: Evidence from a New Test

Although many economists intuitively consider long-run relative PPP to be true, the vast empirical literature has not yet provided conclusive evidence. Chapter 4 re-examines the PPP hypothesis, so as to resolve this contradiction.

We first show that including the deviation from PPP in the regime-switching probabilities of the regime-switching mean model of Chapter 3 yields a new test approach for PPP. This approach is particularly appealing, because PPP can now be tested within a long swings framework and the results of Chapter 3 suggest that long swings are a systematic part of the exchange rate generating process.

The main empirical contribution of Chapter 4 is that we find evidence in favor of PPP for the world's three main U.S. dollar exchange rates over the post-Bretton-Woods period. Hence, in combination with the findings of the previous chapter, we conclude that long swings in exchange rates lead to deviations from PPP in the short-run, but that in the long-run the equilibrating forces behind PPP govern the swings in such a way that the PPP hypothesis holds.

Although the validity of PPP is interesting in itself, it is also relevant for other research. For instance, PPP is a building block of many traditional structural exchange rate models, so that its validity underscores their usefulness for long-run exchange rate determination. Moreover, our result helps the development of new exchange rate models. As an example, consider the theory of exchange rate bubbles, where the exchange rate can diverge from its equilibrium value as determined by an economic model (see De Grauwe (1990)). The dependence of the long swings on the PPP deviation in our model indicates that bubbles tend to burst when the deviation from PPP becomes large. Hence, including PPP in bubble theories seems fruitful.

Chapter 4 also presents support for the commonly held view that goods arbitrage is one of the factors underlying PPP. Moreover, the existence of long-run PPP improves predictions of the direction of exchange rate changes compared to the random walk and the long swings model of Chapter 3. This holds especially for long horizons, presumably because PPP is a long-run phenomenon.

The forecasting exercise also shows that the advantage of incorporating PPP for point prediction is not yet clear. The post-Bretton-Woods data period is too short compared to the length of the swings to get sufficiently accurate in-sample estimates of the regime-switching parameters. Obtaining good forecasts of regime switches seems crucial for the forecast quality of regime-switching models. It may help to pool several exchange rate series in a panel data set and then impose some cross-sectional parameter restrictions to increase estimation accuracy. Also the inclusion of other variables, as suggested in the concluding discussion of Chapter 3, may improve the usefulness of regime-switching models for forecasting.

Another issue for future research concerns tests for other long-run relationships. For instance, think of the long-run quantity theory of money, stating that the price level is proportional to the money supply in the long term. Regime-switching models can presumably provide a framework for such tests as well. In general, regime-switching models may offer an alternative for unit-root tests that are commonly employed to test for long-run relations.

Chapter 5: Have Exchange Rates Become More Closely Tied? Evidence from a New Multivariate GARCH Model

In this chapter we analyze exchange rate correlations. Hence, this chapter is complementary to Chapter 2 on the variance and Chapters 3 and 4 on the mean of exchange rates.

Since economic intuition tells us that exchange rate correlations change over time, we need a model that allows for that. Therefore, we introduce a new multivariate GARCH model. Our model is easy to estimate, since one only has to estimate a univariate GARCH model for each principal component of the exchange rates. Nevertheless, we show in a stylized example that the model has quite realistic economic implications for exchange rate correlations.

The empirical results on eight U.S. dollar exchange rates show that the model provides a better fit than two popular existing multivariate GARCH models.

First, it outperforms the constant conditional correlations model of Bollerslev (1990). Second, our model generalizes usual factor GARCH models (see Ng, Engle and Rothschild (1992)) by taking as many factors as exchange rates. This appears important from an empirical point of view, as the factors neglected in usual factor GARCH contain important information for exchange rate variances and correlations.

Given the appealing theoretical and empirical characteristics, we use our model to analyze the correlations between the dollar exchange rates over time. We find a pattern that is in line with economic intuition and which shows that exchange rates have become less instead of more closely tied since the eighties. In particular, the EMS crisis in 1992 lowered correlations in the nineties after a long period of high correlations in the eighties. The effect of the advent of EMU may be reflected in the rise of the correlation between the mark-dollar and the lira-dollar rates in the second half of the nineties, particularly because the mark-dollar correlation with the pound-dollar rate continued to fall.

So far, we have concentrated on GARCH in a multivariate setting. However, it is important to realize that our indirect approach via the principal components is not restricted to GARCH. In fact, any univariate model for the principal components can be used to derive a practical multivariate model. This offers a wide range of applications of our approach. For instance, when analyzing stock or bond return correlations, one can take account of asymmetric volatilities, GARCH-in-mean effects and other deviations from standard GARCH (see Bollerslev et al. (1992)). Furthermore, our approach can form the basis for multivariate extensions of other volatility models, such as stochastic volatility, regime-switching GARCH and fractionally integrated GARCH.

One can even combine the principal components approach with the regime-switching variance and regime-switching mean models of the previous chapters. This would require a joint process for the mean and variance regimes. Since the mean and variance of exchange rates may well be independent (for instance, see Engle, Ito and Lin (1990)), assuming that the joint regime probability is simply the product of the mean and variance regime probabilities seems reasonable. Hence, it is practically feasible to build a multivariate model with both regime switches in the mean and the variance.

Chapter 6: Why is it so Difficult to Find an Effect of Exchange Rate Risk on Trade?

Chapter 6 concerns the effect of exchange rate risk on trade. The common opinion is that exchange risk affects trade negatively. However, empirical studies, mainly time series papers, produce ambiguous results. In this chapter we analyze the reason for that.

We first develop a trade model, so that we can estimate the effect of risk on trade for our specific data on monthly bilateral aggregate U.S. exports to the G-7 countries. The model contains two methodological contributions to the literature. First, it uses a multi-months-ahead exchange rate risk measure that is more accurate than existing measures, mainly because of the use of daily exchange rates. Second, we pay special attention to the dynamic structure of the model by introducing a new lag structure for distributed lag models. It uses the Poisson probability function to distribute the total effect of a regressor over time.

The empirical results demonstrate that, as expected, foreign income affects U.S. exports positively and rather quickly, with the largest effect occurring after about one quarter. Exports react much slower to changes in the real exchange rate distribution, as the maximal effect happens only after about one year. The expected real exchange rate level has the normal positive effect, but real exchange rate risk has no clear effect.

To explain this latter, commonly reported finding, we examine the long-term (about one year) risk that is relevant for goods traders in more detail. Long-term risk appears rather constant over time with only short-term deviations from average risk. In our opinion, this is the reason why it is so difficult to find an effect of exchange rate risk on trade from time series data.

It is important to realize that our conclusion concerns countries with low time-variation in long-term real exchange rate risk, such as most developed countries. It would be interesting to study the effect of risk on trade flows between countries with more time-variation in risk, for instance, developing countries. In addition, employing cross-sectional variation in exchange risk may prove fruitful, either in a panel setup or in a pure cross-sectional analysis. This is left for future research.

In total, this book provides empirical answers to several significant economic questions. These answers are based on modern econometric methods, either refinements of existing techniques or newly developed approaches. Both the economic results and the econometric novelties embody contributions of this thesis to the existing literature on exchange rates and international trade.

References

Abuaf, N. and P. Jorion (1990), "Purchasing Power Parity in the Long Run," *Journal of Finance*, 45, 157-174.

Almekinders, G.J. and S.C.W. Eijffinger (1991), "Empirical Evidence on Foreign Exchange Market Intervention: Where Do We Stand?" *Weltwirtschaftliches Archiv*, 127, 645-677.

Andersen, T.G. and T. Bollerslev (1998), "Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts," *International Economic Review*, 39, 885-905.

Andersen, T.G., T. Bollerslev, F.X. Diebold and P. Labys (1999), "The Distribution of Exchange Rate Volatility," NBER Working Paper No. 6961.

Ang, A. and G. Bekaert (1998), "Regime Switches in Interest Rates," NBER Working Paper No. 6508.

Arize, A.C. (1995), "The Effects of Exchange Rate Volatility on U.S. Exports: An Empirical Investigation," *Southern Economic Journal*, 62, 34-43.

Bacchetta, P. and E. van Wincoop (1998), "Does Exchange Rate Stability Increase Trade and Capital Flows?" NBER Working Paper No. 6704.

Baillie, R.T., T. Bollerslev and H.O. Mikkelsen (1996), "Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 74, 3-30.

Bank for International Settlements (1999), *Central Bank Survey of Foreign Exchange and Derivatives Market Activity 1998*, Basle.

Bartholomew, D.J. (1987), *Latent Variable Models and Factor Analysis*, London: Charles Griffin.

Bayoumi, T. and B. Eichengreen (1998), "Exchange Rate Volatility and Intervention: Implications of the Theory of Optimum Currency Areas," *Journal of International Economics*, 45, 191-209.

Bertero, E. and C. Mayer (1990), "Structure and Performance: Global Interdependence of Stock Markets around the Crash of October 1987," *European Economic Review*, 34, 1155-1180.

Bini-Smaghi, L. (1991), "Exchange Rate Variability and Trade: Why is it so Difficult to Find Any Empirical Relationship?" *Applied Economics*, 23, 927-936.

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307-327.

Bollerslev, T. (1990), "Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate GARCH Approach," *Review of Economics and Statistics*, 72, 498-505.

Bollerslev, T., R.Y. Chou and K.F. Kroner (1992), "ARCH Modeling in Finance," *Journal of Econometrics*, 52, 5-59.

Bollerslev, T., R.F. Engle and J.M. Wooldridge (1988), "A Capital Asset Pricing Model with Time-Varying Covariances," *Journal of Political Economy*, 96, 116-131.

Broll, U. and B. Eckwert (1997), "Exchange Rate Volatility and International Trade," University of Munich, Working Paper No. 9715.

Cai, J. (1994), "A Markov Model of Unconditional Variance in ARCH," *Journal of Business and Economic Statistics*, 12, 309-316.

Caporale, T. and K. Doroodian (1994), "Exchange Rate Variability and the Flow of International Trade," *Economics Letters*, 46, 49-54.

Choi, I. and B.S. Chung (1995), "Sampling Frequency and the Power of Tests for a Unit Root: A Simulation Study," *Economics Letters*, 49, 131-136.

Chowdury, A.R. (1993), "Does Exchange Rate Volatility Depress Trade Flows? Evidence from Error-Correction Models," *Review of Economics and Statistics*, 75, 700-706.

Confederation of British Industry (1989), *European Monetary Union: A Business Perspective*, London.

Côté, A. (1994), "Exchange Rate Volatility and Trade - a Survey," Bank of Canada, Working Paper No. 945.

Cushman, D.O. (1983), "The Effects of Real Exchange Rate Risk on International Trade," *Journal of International Economics*, 15, 45-63.

Cushman, D.O. (1986), "Has Exchange Risk Depressed International Trade? The Impact of Third-Country Exchange Risk," *Journal of International Money and Finance*, 5, 361-379.

Darbar, S.M. and P. Deb (1997), "Co-Movements in International Equity Markets," *Journal of Financial Research*, 20, 305-322.

De Grauwe, P. (1990), *International Money: Post-War Trends and Theories*, Oxford: Oxford University Press.

Dewachter, H. (1997), "Sign Predictions of Exchange Rate Changes: Charts as Proxies for Bayesian Inferences," Weltwirtschaftliches Archiv, 133, 39-55.

Diebold, F.X. and J.A. Nason (1990), "Nonparametric Exchange Rate Prediction?" *Journal of International Economics*, 28, 315-332.

Diebold, F.X. and M. Nerlove (1989), "The Dynamics of Exchange Rates Volatility: A Multivariate Latent Factor ARCH Model," *Journal of Applied Econometrics*, 4, 1-21.

Durland, J.M. and T.H. McCurdy (1994), "Duration-Dependent Transitions in a Markov Model of U.S. GNP Growth," *Journal of Business & Economic Statistics*, 12, 279-288.

Edison, H.J. (1993), "The Effectiveness of Central-Bank Intervention: A Survey of the Literature after 1982," Princeton University, *Special Papers in International Economics*, 18.

Engel, C. (1994), "Can the Markov Switching Model Forecast Exchange Rates?" *Journal of International Economics*, 36, 151-165.

Engel, C. and J.D. Hamilton (1990), "Long Swings in the Dollar: Are They in the Data and Do Markets Know It?" *American Economic Review*, 80, 689-713.

Engle, R.F. (1982), "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation," *Econometrica*, 50, 987-1008.

Engle, R.F. and T. Bollerslev (1986), "Modeling the Persistence of Conditional Variances," *Econometric Reviews*, 5, 1-50.

Engle, R.F., T. Ito and W.-L. Lin (1990), "Meteor Showers or Heat Waves? Heteroskedastic Intra-daily Volatility in the Foreign Exchange Market," *Econometrica*, 58, 525-542.

Engle, R.G. and K.F. Kroner (1995), "Multivariate Simultaneous Generalized ARCH," *Econometric Theory*, 11, 122-150.

Engle, R.F., V.K. Ng and M. Rothschild (1990), "Asset Pricing with a Factor-ARCH Covariance Structure: Empirical Estimates for Treasury Bills," *Journal of Econometrics*, 45, 213-237.

EU Commission (1990), "One Market, One Money," *European Economy*, 44.

Filardo, A.J. (1994), "Business-Cycle Phases and Their Transitional Dynamics," *Journal of Business & Economic Statistics*, 12, 299-308.

Fiorentini, G., E. Sentana and N. Shephard (1998), "Exact Likelihood-Based Estimation of Conditionally Heteroskedastic Factor Models," Paper presented at Econometric Society European Meeting 98, Berlin.

Fountas, S. and K. Aristotelous (1999), "Has the European Monetary System Led to More Exports? Evidence from Four European Union Countries," *Economics Letters*, 62, 357-363.

Frankel, J.A. (1986), "International Capital Mobility and Crowding-out in the U.S. Economy: Imperfect Integration of Financial Markets or Goods Markets?" in: R.W. Hafer (ed.), *How Open is the U.S. Economy?* Lexington: Lexington Books, 33-67.

Frankel, J.A. (1990), "Zen and the Art of Modern Macroeconomics: A Commentary," in: W.S. Haraf and T.D. Willett (eds.), *Monetary Policy for a Volatile Global Economy*, Washington DC: American Enterprise Institute for Public Policy Research, 117-123.

Frankel, J.A. and A.K. Rose (1996), "A Panel Project on Purchasing Power Parity: Mean Reversion Within and Between Countries," *Journal of International Economics*, 40, 209-224.

Froot, K.A. and K. Rogoff (1996), "Perspectives on PPP and Long-Run Exchange Rates," in: G. Grossman and K. Rogoff (eds.), *Handbook of International Economics*, 3, 1647-1688.

Gagnon, J.E. (1993), "Exchange Rate Variability and the Level of International Trade," *Journal of International Economics*, 34, 269-287.

Gallant, A.R. (1987), *Nonlinear Statistical Models*, New York: Wiley & Sons.

Garcia, R. (1998), "Asymptotic Null Distribution of the Likelihood Ratio Test in Markov Switching Models," *International Economic Review*, 39, 763-788.

Garcia, R. and P. Perron (1996), "An Analysis of the Real Interest Rate under Regime Shifts," *Review of Economics and Statistics*, 78, 111-125.

Ghysels, E. (1994), "On the Periodic Structure of the Business Cycle," *Journal of Business & Economic Statistics*, 12, 289-298.

Ghysels, E., A. Harvey and E. Renault (1996), "Stochastic Volatility," in: Maddala, G.S. and C.R. Rao (eds.), *Handbook of Statistics*, 14, Amsterdam: North Holland, 119-191.

Goldstein, M. and M.S. Khan (1985), "Income and Price Effects in Foreign Trade," in: R.W. Jones and P.B. Kenen (eds.), *Handbook of International Economics*, Amsterdam: North-Holland, 1041-1105.

Goodwin, T.H. (1993), "Business-Cycle Analysis With a Markov-Switching Model," *Journal of Business & Economic Statistics*, 11, 331-339.

Gray, S.F. (1996a), "Modeling the Conditional Distribution of Interest Rates as a Regime-Switching Process," *Journal of Financial Economics*, 42, 27-62.

Gray, S.F. (1996b), "An Analysis of Conditional Regime-Switching Models," Working Paper, Fuqua School of Business, Duke University.

Hamilton, J.D. (1989), "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle", *Econometrica*, 57, 357-384.

Hamilton, J.D. (1994), *Time Series Analysis*, Princeton: Princeton University Press.

Hamilton, J.D. and R. Susmel (1994), "Autoregressive Conditional Heteroskedasticity and Changes in Regime," *Journal of Econometrics*, 64, 307-333.

Hansen, B.E. (1992), "The Likelihood Ratio Test under Nonstandard Conditions: Testing the Markov Switching Model of GNP," *Journal of Applied Econometrics*, 7, S61-S82.

Härdle, W. and O. Linton (1994), "Applied Nonparametric Methods," in: R.F. Engle and D.L. McFadden (eds.), *Handbook of Econometrics*, 4, Amsterdam: North-Holland, 2297-2341.

Kaminsky, G. (1993), "Is There a Peso Problem? Evidence from the Dollar/Pound Exchange Rate, 1976-1987," *American Economic Review*, 83, 450-472.

King, M., E. Sentana and S. Wadhwani (1994), "Volatility and Links between National Stock Markets," *Econometrica*, 62, 901-933.

Klein, M.W. (1990), "Sectoral Effects of Exchange Rate Volatility on United States Exports," *Journal of International Money and Finance*, 9, 299-308.

Koch, P.D. and T.W. Koch (1991), "Evolution in Dynamic Linkages across National Stock Indexes," *Journal of International Money and Finance*, 10, 231-251.

Kodde, D.A. and F.C. Palm (1986), "Wald Criteria for Jointly Testing Equality and Inequality Restrictions," *Econometrica*, 54, 1243-1248.

Koedijk, K.G., P.C. Schotman and M.A. van Dijk (1998), "The Re-emergence of PPP in the 1990s," *Journal of International Money and Finance*, 19, 51-61.

Kroner, K.F. and W.D. Lastrapes (1993), "The Impact of Exchange Rate Volatility on International Trade: Reduced Form Estimates Using the GARCH-in-Mean Model," *Journal of International Money and Finance*, 12, 298-318.

Krugman, P.R. and M. Obstfeld (1991), *International Economics: Theory and Policy*, New York: Harper Collins.

Lam, P. (1990), "The Hamilton Model with a General Autoregressive Component," *Journal of Monetary Economics*, 26, 409-432.

Lamoureux, C.G. and W.D. Lastrapes (1990), "Persistence in Variance, Structural Change, and the GARCH Model," *Journal of Business and Economic Statistics*, 8, 225-234.

Lin, W.-L. (1992), "Alternative Estimators for Factor GARCH Models - A Monte Carlo Comparison," *Journal of Applied Econometrics*, 7, 259-279.

Longin, F. and B. Solnik (1995), "Is the Correlation in International Equity Returns Constant: 1960-1990?" *Journal of International Money and Finance*, 14, 3-26.

Loopesko, B.E. (1984), "Relationships among Exchange Rates, Intervention and Interest Rates: An Empirical Investigation," *Journal of International Money and Finance*, 3, 257-277.

Lucas, R.E. (1976), "Econometric Policy Evaluation: A Critique," *Carnegie-Rochester Series on Public Policy*, 1, 19-46.

MacDonald, R. and M.P. Taylor (1992), "Exchange Rate Economics: A Survey," *IMF Staff Papers*, 39, 1-57.

Mark, N.C. (1990), "Real and Nominal Exchange Rates in the Long Run; An Empirical Investigation," *Journal of International Economics*, 28, 115-136.

Mark, N.C. and D.-Y. Choi (1997), "Real Exchange Rate Prediction over Long Horizons," *Journal of International Economics*, 43, 29-60.

Marquez, J. (1999), "Long-Period Trade Elasticities for Canada, Japan, and the United States," *Review of International Economics*, 7, 102-116.

Meese, R.A. and K. Rogoff (1983), "Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample?" *Journal of International Economics*, 14, 3-24.

Meese, R.A. and K. Rogoff (1988), "Was it Real? The Exchange Rate Interest Differential Relation over the Modern Floating Exchange Rate Period," *Journal of Finance*, 43, 933-948.

Merton, R.C. (1980), "On Estimating the Expected Return on the Market: An Exploratory Investigation," *Journal of Financial Economics*, 8, 323-361.

Mussa, M. (1986), "Nominal Exchange Rate Regimes and the Behavior of the Real Exchange Rate," in: K. Brunner and A.H. Meltzer (eds.), *Real Business Cycles, Real Exchange Rates and Actual Policies*, New York: North Holland, 117-213.

Newey, W.K. and K.D. West (1987), "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703-708.

Newey, W.K. and K.D. West (1994), "Automatic Lag Selection in Covariance Matrix Estimation," *Review of Economic Studies*, 61, 631-653.

Ng, V., R.F. Engle and M. Rothschild (1992), "A Multi-Dynamic-Factor Model for Stock Returns," *Journal of Econometrics*, 52, 245-266.

O'Connell, P.G.J. (1998), "The Overvaluation of Purchasing Power Parity," *Journal of International Economics*, 44, 1-19.

Pagan, A.R. and G.W. Schwert (1990), "Alternative Models for Conditional Stock Volatility," *Journal of Econometrics*, 45, 267-290.

Page, S.A.B. (1981), "The Choice of Invoicing Currency in Merchandise Trade," *National Institute Economic Review*, 85, 60-72.

Papell, D. (1997), "Searching for Stationarity: Purchasing Power Parity under the Current Float," *Journal of International Economics*, 43, 313-332.

Perron, P. (1989), "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis," *Econometrica*, 57, 1361-1401.

Pozo, S. (1992), "Conditional Exchange Rate Volatility and the Volume of International Trade: Evidence from the Early 1900s," *Review of Economics and Statistics*, 74, 325-329.

Qian, Y. and P. Varangis (1994), "Does Exchange Rate Volatility Hinder Export Growth?" *Empirical Economics*, 19, 371-396.

Sawyer, W.C. and R.L. Sprinkle (1997), "An Examination of U.S. Imports and Exports of Finished Manufactures," *Rivista Internazionale di Scienze Economiche e Commerciali*, 44, 139-148.

Sentana, E. (1992), "Identification of Multivariate Conditionally Heteroskedastic Factor Models," London School of Economics, Discussion Paper No. 139.

Stokman, A.C.J. (1995), "Terms of Payment in International Trade," Ph. D. Thesis, University of Amsterdam.

Taylor, M.P. and H.L. Allen (1992), "The Use of Technical Analysis in the Foreign Exchange Market," *Journal of International Money and Finance*, 11, 304-314.

Viaene, J.-M. and C.G. de Vries (1992), "International Trade and Exchange Rate Volatility," *European Economic Review*, 36, 1311-1321.

Wei, S.-J. (1999), "Currency Hedging and Goods Trade," *European Economic Review*, 43, 1371-1394.

West, K.D. and D. Cho (1995), "The Predictive Ability of Several Models of Exchange Rate Volatility," *Journal of Econometrics*, 69, 367-391.

White, H. (1980), "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica*, 48, 817-838.

World Trade Organization (1998), *Annual Report*, Geneva.

Samenvatting (Summary in Dutch)

Het belang van wisselkoersen en internationale handel is de laatste decennia enorm gegroeid. Het is daarom niet verwonderlijk, dat beide variabelen een belangrijke rol spelen bij allerlei beleidsdiscussies. Zo kunnen handelsproblemen tengevolge van een overgewaardeerde munt de aanleiding vormen voor valutamarktinterventies. Een ander voorbeeld betreft de Europese monetaire eenwording (EMU). Eén van de belangrijkste economische motivaties daarvoor was de algemene notie dat wisselkoersstabiliteit de handel bevordert.

Vanwege hun belang voor beleidsvraagstukken, is er tot nu toe reeds veel economisch onderzoek naar wisselkoersen en internationale handel gedaan. Desalniettemin heeft men de onderliggende processen nog niet in zijn geheel kunnen doorgronden.

De eerste bijdrage van dit proefschrift is dat we nader inzicht verschaffen in een aantal van die economische processen. Dit gebeurt vanuit een empirische invalshoek, met de Verenigde Staten als centraal land. We onderzoeken, kortgezegd, de volgende vijf onderwerpen (verderop in deze samenvatting volgt een uitgebreidere beschrijving). Ten eerste analyseren we de volatiliteit (beweeglijkheid) van wisselkoersen en introduceren we een manier om deze beter te voorspellen. Daarna toetsen we of wisselkoersen gekenmerkt worden door lange swings. Deze kunnen veroorzaakt worden door veranderingen in economisch beleid. Zo is de overgang naar een anti-inflatie beleid door Volcker in de V.S. in 1979 wellicht de reden geweest voor de sterke appreciatie van de Amerikaanse dollar in de eerste helft van de jaren tachtig. Ten derde onderzoeken we de geldigheid van de koopkrachtpariteit (*purchasing power parity*, PPP) hypothese, een belangrijk element in veel economische theorieën: is de wisselkoers op lange termijn proportioneel aan de verhouding van de prijsniveaus van beide landen? Daarna bekijken we de correlaties tussen wisselkoersen, met name hun ontwikkelingen in de tijd. Hoe heeft bijvoorbeeld de wisselkoerscoördinatie binnen het Europese Monetaire Systeem (EMS) deze correlaties beïnvloed? Als vijfde en laatste onderwerp van dit proefschrift onderzoeken we het effect van wisselkoersrisico op handel. Dit effect speelt een belangrijke rol bij de keuze tussen een systeem van vaste of flexibele wisselkoersen en bij het interveniëren op valutamarkten om wisselkoersen te stabiliseren.

De tweede bijdrage van het boek is econometrisch van aard. Bij het analyseren van de zojuist genoemde onderwerpen blijkt, dat bestaande econometrische technieken vaak niet direct toepasbaar zijn. Daarom verfijnen we diverse bestaande technieken en ontwikkelen waar nodig nieuwe. Deze methodologische bijdragen zijn nuttig bij toekomstig onderzoek. Het betreft, kortgezegd, onder andere de volgende technieken (een uitgebreidere beschrijving volgt later). We introduceren een nieuwe manier om Markov *regime-switches* (toestandsveranderingen) in de variantie te combineren met gegeneraliseerde autoregressieve conditionele variantie (GARCH) processen om de veelgebruikte standaard GARCH volatiliteitsvoorspellingen te verbeteren. We gebruiken modellen met *regime-switches* in het niveau in plaats van de variantie om de lange swings in wisselkoersen te beschrijven en ook om PPP te toetsen. We introduceren een nieuw multivariaat GARCH model om de ontwikkeling van wisselkoerscorrelaties in de tijd te modelleren. Tenslotte laten we zien hoe wisselkoersen op dagbasis gebruikt kunnen worden om meetfouten in lange-termijn risicomaten te verkleinen.

Het vervolg van deze samenvatting geeft een meer gedetailleerde beschrijving van de twee genoemde bijdragen van dit proefschrift. Voor elk hoofdstuk, afgezien van het inleidende hoofdstuk 1 en het concluderende hoofdstuk 7, beschrijven we de economische probleemstelling, de idee achter de econometrische aanpak, de belangrijkste bijdragen aan de bestaande literatuur en het nut voor toekomstig onderzoek.

Hoofdstuk 2: Het Verbeteren van GARCH Volatiliteitsvoorspellingen met een Nieuw *Regime-Switching* GARCH Model

Hoofdstuk 2 gaat over het voorspellen van de volatiliteit van wisselkoersen. Veel onderzoekers gebruiken GARCH modellen daarvoor, omdat die modellen het empirisch feit kunnen verklaren dat wisselkoersschokken vaak enige tijd doorwerken in de volatiliteit. Wij laten echter zien dat de GARCH voorspellingen te variabel zijn.

Om dit probleem te verhelpen, generaliseren we het GARCH model door het toevoegen van een extra mechanisme om de persistentie van wisselkoersschokken te verklaren. Dit doen we door het onderscheiden van twee *regimes* (toestanden) met verschillende niveaus van volatiliteit. Persistentie van deze *regimes* verklaart al een deel van de persistentie van schokken; het overgebleven deel wordt gemodelleerd door een apart GARCH proces binnen elk *regime*. Het resulterende model noemen we het *regime-switching* GARCH model.

De belangrijkste bijdrage van hoofdstuk 2 betreft de manier waarop we het *regime-switching* aspect met het GARCH model combineren. Ten opzichte van bestaande studies, zoals Gray (1996a), heeft onze aanpak het voordeel dat het voorspellen van de volatiliteit meerdere perioden vooruit veel eenvoudiger is.

Hierdoor kunnen wij de standaard GARCH voorspellingen empirisch vergelijken met die van *regime-switching* GARCH. Gebruikmakend van gegevens over de drie belangrijkste dollar wisselkoersen (Duitse mark, Japanse yen en Britse pond), zien we dat *regime-switching* GARCH voorspellingen niet te variabel zijn (in tegenstelling tot standaard GARCH) en dat ze beter zijn dan de standaard GARCH voorspellingen.

De theoretische en empirische bijdragen zoals hierboven beschreven kunnen beide nuttig zijn voor toekomstig onderzoek. De voorgestelde combinatie van GARCH en *regimes* voor de volatiliteit van een variabele is ook toepasbaar in modellen met GARCH en *regimes* voor de variabele zelf, zoals in de hoofdstukken 3 en 4 blijkt. De verbeteringen in de kwaliteit van de volatiliteitsvoorspellingen is bijvoorbeeld nuttig bij analyses van het effect van volatiliteit op andere financiële grootheden, zoals optieprijzen.

Hoofdstuk 3: Worden Wisselkoersen Werkelijk Gekenmerkt door Lange Swings?

In de literatuur is veel aandacht besteed aan het modelleren van het niveau van wisselkoersen. Er zijn veel structurele wisselkoersmodellen ontwikkeld, maar met name op korte termijn wordt hun empirische validiteit vaak betwijfeld. Daarom gebruiken onderzoekers vaak een *random walk* model voor de wisselkoers.

Vanuit economisch oogpunt is de *random walk* echter problematisch. Zo houdt het geen rekening met effecten van veranderingen in economisch beleid op het wisselkoersproces. Het is bijvoorbeeld best mogelijk, dat de overgang naar een krap monetair beleid in de V.S. onder Volcker in 1979 heeft geleid tot een verhoging van de structurele dollarappreciatie, waardoor de waarde van de dollar in de eerste helft van de jaren tachtig zo scherp is gestegen. De Plaza-interventie van 1985 om de dollar goedkoper te maken heeft wellicht de structurele appreciatie omgezet in een depreciatie, waardoor de eerdere dollarstijging teniet werd gedaan.

Beide voorbeelden laten zien dat beleidsaanpassingen effect kunnen hebben op de trend van wisselkoersen en derhalve lange swings kunnen veroorzaken. De kernvraag van hoofdstuk 3 is of dergelijke lange swings werkelijk een structureel onderdeel van het wisselkoersproces zijn.

Om de lange swings te modelleren maken we wederom gebruik van *regimes*. In tegenstelling tot het *regime-switching* model van hoofdstuk 2, gebruiken we nu echter de regimes voor de wisselkoersverandering zelf, niet voor de variantie ervan. Zo kan de persistentie van een appreciatie en daarna een depreciatie *regime* de dollarswing in de jaren tachtig verklaren. Om het bestaan van lange swings na te gaan, toetsen we de *random walk* versus het *regime-switching* model.

Op basis van kwartaaldata voor de drie belangrijkste dollarkoersen vinden we geen aanwijzingen voor lange swings. Gebruikmakend van weekdata voor dezelfde koersen blijkt echter dat lange swings wel bestaan. Kennelijk kan de datafrequentie van belang zijn bij het toetsen op lange-termijn eigenschappen.

Onze conclusie dat wisselkoersen gekenmerkt worden door lange swings kan duiden op de relevantie van economische beleidswijzigingen voor wisselkoersprocessen. Het is daarom interessant om na te gaan of bijvoorbeeld aankondigingen van beleidsveranderingen of valutamarktinterventies informatie over toekomstige wisselkoersontwikkelingen bevatten. Hoofdstuk 3 laat zien hoe het *regime-switching* model daarbij van dienst kan zijn.

Hoofdstuk 4: Aanwijzingen voor Koopkrachtpariteit op Basis van een Nieuwe Toets

Koopkrachtpariteit is een van de belangrijkste theorieën in de internationale economie. De meeste economen gaan er vanuit dat koopkrachtpariteit (PPP) geldt. De bestaande empirische literatuur trekt dat echter in twijfel. In hoofdstuk 4 toetsen we de geldigheid van PPP opnieuw, echter op een nieuwe manier.

De voorgestelde toets op PPP is ingebed in het *regime-switching* model van hoofdstuk 3. Dit heeft als voordeel, dat we automatisch rekening houden met het bestaan van lange swings in wisselkoersen, zoals aangetoond in het vorige hoofdstuk. De idee van de toets is dat PPP geldt als de lange swings zich rond de PPP-koers (de verhouding van de prijsniveaus van de landen) bewegen.

De belangrijkste empirische bijdrage van het hoofdstuk is, dat de nieuwe toets aangeeft dat PPP geldt voor de drie belangrijkste dollarkoersen. Tezamen met het resultaat van hoofdstuk 3 vinden we dus, dat de lange swings leiden tot afwijkingen van PPP op korte termijn, maar dat evenwichtsherstellende krachten ervoor zorgen dat PPP op lange termijn wel geldt.

De tweede bijdrage van hoofdstuk 4 betreft het mechanisme waardoor het PPP-evenwicht op lange termijn hersteld wordt. Men beweert vaak dat evenwichtsherstel veroorzaakt wordt door arbitrage op goederenmarkten. Deze goederenmarkten zijn meer geïntegreerd dan vroeger, waardoor arbitrage nu eenvoudiger is. Omdat we vinden dat de duur van PPP-afwijkingen is afgenomen in de loop van de tijd, ondersteunen de data dat goederenmarktarbitrage inderdaad bijdraagt aan het herstel van PPP-evenwicht.

De geldigheid van PPP is van belang voor toekomstig onderzoek. Omdat PPP een onderdeel is van vele traditionele wisselkoerstheorieën, ondersteunt ons resultaat de zin van het gebruik van die theorieën voor wisselkoersanalyses op lange termijn. Tevens is het zinvol om PPP te incorporeren in nieuwe wisselkoerstheorieën. Zo geeft de analyse van hoofdstuk 4 aan, dat er een grotere kans bestaat op het uiteenspatten van een wisselkoersbubbel als de afwijking van PPP groter wordt. Het is dus nuttig om rekening te houden met PPP in de recente bubbeltheorieën voor wisselkoersen.

Hoofdstuk 5: Zijn Wisselkoersen Nauwer aan Elkaar Gerelateerd dan Vroeger? Resultaten van een Nieuw Multivariaat GARCH Model

In hoofdstuk 5 onderzoeken we correlaties tussen wisselkoersen. Dit hoofdstuk is dus complementair aan hoofdstuk 2 over de variantie en de hoofdstukken 3 en 4 over het niveau van wisselkoersen.

Volgens de economische intuïtie veranderen wisselkoerscorrelaties in de tijd. Bijvoorbeeld, als het Verenigd Koninkrijk toetreedt tot het wisselkoersmechanisme van het EMS, dan zal de correlatie tussen de pond-dollar en de mark-dollar koersen toenemen. De bestaande literatuur laat echter zien, dat het niet eenvoudig is een praktisch model te ontwikkelen waarin de correlaties mogen variëren over de tijd.

De belangrijkste bijdrage van hoofdstuk 5 is, dat het een nieuw multivariaat GARCH model met tijds variërende correlaties introduceert dat toch eenvoudig te schatten is. Men hoeft alleen (univariate) GARCH modellen te schatten voor elke principale component van de wisselkoersveranderingen. Desalniettemin heeft het model economisch gezien realistische implicaties voor wisselkoerscorrelaties.

De empirische resultaten voor acht dollar wisselkoersen laten zien, dat ons model de data beter beschrijft dan het veelgebruikte constante-conditionele-correlaties model van Bollerslev (1990) en het standaard factor-GARCH model (zie Ng, Engle en Rothschild (1992)). Bovendien is het geschatte verloop van de correlaties in de tijd goed te verklaren op basis van de economische intuïtie. In het bijzonder vinden we, dat de correlaties gedaald zijn sinds de jaren tachtig. Met name de EMS-crisis in 1992 is daar debet aan. De nadering van de EMU is er wellicht de oorzaak van, dat de correlatie tussen de mark-dollar en lire-dollar koersen in de tweede helft van de negentiger jaren is gestegen, temeer omdat de mark-dollar correlatie met de pond-dollar koers is blijven dalen.

Voor toekomstig onderzoek is het belangrijk te beseffen, dat onze indirecte aanpak via principale componenten niet alleen nuttig is voor multivariate GARCH modellen. Men kan elk univariaat model voor de principale componenten gebruiken om een praktisch multivariaat model te verkrijgen. Zo kan men bijvoorbeeld multivariate modellen met stochastische volatiliteit, *regime-switching* en fractioneel geïntegreerd GARCH ontwikkelen.

Hoofdstuk 6: Waarom is het zo Moeilijk om een Effect van Wisselkoersrisico op Handel te Vinden?

Men gaat er in het algemeen van uit, dat wisselkoersrisico een negatief effect op handel heeft. De bestaande empirische literatuur, voornamelijk gebaseerd op tijdreeksanalyses, levert echter geen duidelijk effect. In hoofdstuk 6 proberen we de reden daarvoor te achterhalen.

Allereerst schatten we een handelsmodel voor onze specifieke data, namelijk betreffende de bilaterale geaggregeerde export van de V.S. naar de overige G7-landen. We schenken daarbij speciale aandacht aan het meten van wisselkoersrisico en aan de dynamische specificatie van het model. We vinden dat, zoals verwacht, een toename van het buitenlands inkomen de Amerikaanse export bevordert en dat het grootste effect na ongeveer een kwartaal optreedt. De export reageert veel trager op de wisselkoers, omdat het maximale effect daarvan pas optreedt na ongeveer een jaar. De verwachte wisselkoers heeft het normale positieve effect. Het wisselkoersrisico heeft, net als in de literatuur, geen duidelijk effect op handel.

Om dit laatste resultaat te verklaren, bekijken we het lange-termijn (ongeveer een jaar) risico dat relevant is voor goederenhandelaren. Dit risico blijkt vrij constant in de tijd te zijn, met slechts kortdurende afwijkingen van het gemiddelde risico. Volgens ons is dit de reden waarom het zo moeilijk is om een effect van wisselkoersrisico op handel te vinden met tijdreeksgegevens. Dit inzicht geeft de belangrijkste bijdrage van hoofdstuk 6 aan de literatuur.

Het hoofdstuk geeft derhalve aan, dat bij het onderzoeken van de effecten van wisselkoersrisico op handel tijdreeksstudies vaak niet erg zinvol zijn. Het is nuttiger om meer gebruik te maken van de verschillen in risico's tussen wisselkoersen onderling. Met andere woorden, cross-sectionele of panel studies liggen meer voor de hand.

In zijn geheel bezien, levert deze dissertatie empirische antwoorden op enkele significante economische vragen. Deze antwoorden zijn gebaseerd op moderne econometrische methoden, hetzij verfijningen van bestaande technieken, hetzij nieuw ontwikkelde methoden. Zowel de economische resultaten als de econometrische vernieuwingen belichamen de bijdragen van dit proefschrift aan de bestaande literatuur over wisselkoersen en internationale handel.

Center for Economic Research, Tilburg University, The Netherlands
Dissertation Series

No.	Author	Title
1	P.J.J. Herings	Static and Dynamic Aspects of General Disequilibrium Theory; ISBN 90 5668 001 3
2*	Erwin van der Krabben	Urban Dynamics: A Real Estate Perspective - An institutional analysis of the production of the built environment; ISBN 90 5170 390 2
3	Arjan Lejour	Integrating or Desintegrating Welfare States? - a qualitative study to the consequences of economic integration on social insurance; ISBN 90 5668 003 X
4	Bas J.M. Werker	Statistical Methods in Financial Econometrics; ISBN 90 5668 002 1
5	Rudy Douven	Policy Coordination and Convergence in the EU; ISBN 90 5668 004 8
6	Arie J.T.M. Weeren	Coordination in Hierarchical Control; ISBN 90 5668 006 4
7	Herbert Hamers	Sequencing and Delivery Situations: a Game Theoretic Approach; ISBN 90 5668 005 6
8	Annemarie ter Veer	Strategic Decision Making in Politics; ISBN 90 5668 007 2
9	Zaifu Yang	Simplicial Fixed Point Algorithms and Applications; ISBN 90 5668 008 0
10	William Verkooijen	Neural Networks in Economic Modelling - An Empirical Study; ISBN 90 5668 010 2
11	Henny Romijn	Acquisition of Technological Capability in Small Firms in Developing Countries; ISBN 90 5668 009 9
12	W.B. van den Hout	The Power-Series Algorithm - A Numerical Approach to Markov Processes; ISBN 90 5668 011 0
13	Paul W.J. de Bijl	Essays in Industrial Organization and Management Strategy; ISBN 90 5668 012 9

* Copies can be ordered from ThelaThesis, Prinseneiland 305, 1013 LP Amsterdam, The Netherlands, phone + 31 20 6255429; fax: +31 20 6203395; e-mail: office@thelathesis.nl

No.	Author	Title
14	Martijn van de Ven	Intergenerational Redistribution in Representative Democracies; ISBN 90 5668 013 7
15	Eline van der Heijden	Altruism, Fairness and Public Pensions: An Investigation of Survey and Experimental Data; ISBN 90 5668 014 5
16	H.M. Webers	Competition in Spatial Location Models; ISBN 90 5668 015 3
17	Jan Bouckaert	Essays in Competition with Product Differentiation and Bargaining in Markets; ISBN 90 5668 016 1
18	Zafar Iqbal	Three-Gap Analysis of Structural Adjustment in Pakistan; ISBN 90 5668 017 X
19	Jimmy Miller	A Treatise on Labour: A Matching-Model Analysis of Labour-Market Programmes; ISBN 90 5668 018 8
20	Edwin van Dam	Graphs with Few Eigenvalues - An interplay between combinatorics and algebra; ISBN 90 5668 019 6
21	Henk Oosterhout	Takeover Barriers: the good, the bad, and the ugly; ISBN 90 5668 020 X
22	Jan Lemmen	Financial Integration in the European Union: Measurement and Determination; ISBN 90 5668 021 8
23	Chris van Raalte	Market Formation and Market Selection; ISBN 90 5668 022 6
24	Bas van Aarle	Essays on Monetary and Fiscal Policy Interaction: Applications to EMU and Eastern Europe; ISBN 90 5668 023 4
25	Francis Y. Kumah	Common Stochastic Trends and Policy Shocks in the Open Economy: Empirical Essays in International Finance and Monetary Policy; ISBN 90 5668 024 2
26	Erik Canton	Economic Growth and Business Cycles; ISBN 90 5668 025 0
27	Jeroen Hoppenbrouwers	Conceptual Modeling and the Lexicon; ISBN 90 5668 027 7
28	Paul Smit	Numerical Analysis of Eigenvalue Algorithms Based on Subspace Iterations; ISBN 90 5668 026 9

No.	Author	Title
29	Uri Gneezy	Essays in Behavioral Economics; ISBN 90 5668 028 5
30	Erwin Charlier	Limited Dependent Variable Models for Panel Data; ISBN 90 5668 029 3
31	Rob Euwals	Empirical Studies on Individual Labour Market Behaviour; ISBN 90 5668 030 7
32	Anurag N. Banerjee	The Sensitivity of Estimates, Inferences, and Forecasts of Linear Models; ISBN 90 5668 031 5
33	Frans A. de Roon	Essays on Testing for Spanning and on Modeling Futures Risk Premia; ISBN 90 5668 032 3
34	Xiangzhu Han	Product Differentiation, Collusion and Standardization; ISBN 90 5668 033 1
35	Marcel Das	On Income Expectations and Other Subjective Data: A Micro-Econometric Analysis; ISBN 90 5668 034 X
36	Jeroen Suijs	Cooperative Decision Making in a Stochastic Environment; ISBN 90 5668 035 8
37	Talitha Feenstra	Environmental Policy Instruments and International Rivalry: A Dynamic Analysis; ISBN 90 5668 036 6
38	Jan Bouwens	The Use of Management Accounting Systems in Functionally Differentiated Organizations; ISBN 90 5668 037 4
39	Stefan Hochguertel	Households' Portfolio Choices; ISBN 90 5668 038 2
40	Henk van Houtum	The Development of Cross-Border Economic Relations; ISBN 90 5668 039 0
41	Jorg Jansen	Service and Inventory Models subject to a Delay-Limit; ISBN 90 5668 040 4
42	F.B.S.L.P. Janssen	Inventory Management Systems: control and information issues; ISBN 90 5668 041 2
43	Henri L.F. de Groot	Economic Growth, Sectoral Structure and Unemployment; ISBN 90 5668 042 0
44	Jenke R. ter Horst	Longitudinal Analysis of Mutual Fund Performance; ISBN 90 5668 043 9

No.	Author	Title
45	Marco Hoeberichts	The Design of Monetary Institutions; ISBN 90 5668 044 7
46	Adriaan Kalwij	Household Consumption, Female Employment and Fertility Decisions: A microeconometric analysis; ISBN 90 5668 045 5
47	Ursula Glunk	Realizing High Performance on Multiple Stakeholder Domains: A Resource-based Analysis of Professional Service Firms in the Netherlands and Germany; ISBN 90 5668 046 3
48	Freek Vermeulen	Shifting Ground: Studies on the Intersection of Organizational Expansion, Internationalization, and Learning; ISBN 90 5668 047 1
49	Haoran Pan	Competitive Pressures on Income Distribution in China; ISBN 90 5668 048 X
50	Laurence van Lent	Incomplete Contracting Theory in Empirical Accounting Research; ISBN 90 5668 049 8
51	Rob Aalbers	On the Implications of Thresholds for Economic Science and Environmental Policy; ISBN 90 5668 050 1
52	Abe de Jong	An Empirical Analysis of Capital Structure Decisions in Dutch Firms; ISBN 90 5668 051 X
53	Trea Aldershof	Female Labor Supply and Housing Decisions; ISBN 90 5668 052 8
54	Jan Fidrmuc	The Political Economy of Reforms in Central and Eastern Europe; ISBN 90 5668 053 6
55	Michael Kosfeld	Individual Decision-Making and Social Interaction; ISBN 90 5668 054 4
56	Aldo de Moor	Empowering Communities - A method for the legitimate user-driven specification of network information systems; ISBN 90 5668 055 2
57	Yohane A. Khamfula	Essays on Exchange Rate Policy in Developing Countries; ISBN 90 5668 056 0
58	Maurice Koster	Cost Sharing in Production Situations and Network Exploitation; ISBN 90 5668 057 9

No.	Author	Title
59	Miguel Rosellon Cifuentes	Essays on Financial Policy, Liquidation Values and Product Markets; ISBN 90 5668 058 7
60	Sharon Schalk	Equilibrium Theory: A salient approach; ISBN 90 5668 059 5
61	Mark Voorneveld	Potential Games and Interactive Decisions with Multiple Criteria; ISBN 90 5668 060 9
62	Edward Droste	Adaptive Behavior in Economic and Social Environments; ISBN 90 5668 061 7
63	Jos Jansen	Essays on Incentives in Regulation and Innovation; ISBN 90 5668 062 5
64	Franc J.G.M. Klaassen	Exchange Rates and their Effects on International Trade; ISBN 90 5668 063 3

CentER Dis

Franc Klaassen received his Master's degrees in econometrics and economics from Tilburg University in 1995. He then followed the Ph.D. program at CentER and the Departments of Econometrics and Economics of the same university. His main research interests are empirical international economics and finance, econometrics and the statistical analysis of sports, in particular tennis.

The role of exchange rates and international trade has grown substantially over the last decades. Therefore, both issues have been important in many policy debates and they have been the focus variables in numerous economic studies. Nevertheless, there is still no full understanding of the processes underlying them. For example, do exchange rates exhibit long swings, possibly caused by infrequent changes in monetary policy? Are exchange rates in the long run governed by economic fundamentals, such as price levels? Does exchange rate stability really benefit trade, as is commonly argued when discussing the desirability of foreign exchange interventions and monetary unification? This dissertation tries to answer such unresolved economic questions. It takes an empirical point of view, using the United States as the central country. The book employs modern econometric methods and, if necessary, introduces new techniques to obtain novel insights about exchange rates and their effects on international trade.

ISBN: 90 5668 063 3